

Symmetry Reduced Dynamics of Charged Molecular Strands

DAVID C. P. ELLIS, FRANÇOIS GAY-BALMAZ, DARRYL D. HOLM,
 VAKHTANG PUTKARADZE & TUDOR S. RATIU

Communicated by M. ORTIZ

Abstract

The equations of motion are derived for the dynamical folding of charged molecular strands (such as DNA) modeled as flexible continuous filamentary distributions of interacting rigid charge conformations. The new feature is that these equations are nonlocal when the screened Coulomb interactions, or Lennard–Jones potentials between pairs of charges, are included. The nonlocal dynamics is derived in the *convective* representation of continuum motion by using modified Euler–Poincaré and Hamilton–Pontryagin variational formulations that illuminate the various approaches within the framework of symmetry reduction of Hamilton’s principle for exact geometric rods. In the absence of nonlocal interactions, the equations recover the classical Kirchhoff theory of elastic rods. The motion equations in the convective representation are shown to arise by a classical Lagrangian reduction associated to the symmetry group of the system. This approach uses the process of affine Euler–Poincaré reduction initially developed for complex fluids. On the Hamiltonian side, the Poisson bracket of the molecular strand is obtained by reduction of the canonical symplectic structure on phase space. A change of variables allows a direct passage from this classical point of view to the covariant formulation in terms of Lagrange–Poincaré equations of field theory. In another revealing perspective, the convective representation of the nonlocal equations of molecular strand motion is transformed into quaternionic form.

Contents

1. Introduction	812
1.1. Physical setup	812
1.2. Mathematical setup	817
1.3. Connection to previous studies	820
2. Motion of exact self-interacting geometric rods	827
2.1. Problem set-up	827

- 2.2. Convective representation of nonlocal potential energy 827
- 2.3. Kinematics 830
- 2.4. Remark on n -dimensional generalization and use of other groups 831
- 3. Derivation of strand dynamics in the convective representation 833
 - 3.1. A modified Euler–Poincaré approach 833
 - 3.2. Conservation laws 838
 - 3.3. Lie–Poisson Hamiltonian structure of strand equations 841
- 4. Introduction to affine Lie group actions 842
- 5. The affine Euler–Poincaré and Lie–Poisson approaches 845
 - 5.1. Notations for semidirect products 845
 - 5.2. Affine Lagrangian and Hamiltonian semidirect product theory 846
 - 5.3. Affine reduction at fixed parameter 849
 - 5.4. Application to the charged strand 856
- 6. New variables: coordinate change and horizontal-vertical split 863
 - 6.1. Motivation in terms of covariant derivatives 864
 - 6.2. The case of charged strands 864
 - 6.3. Change of coordinates 865
 - 6.4. The general case 869
- 7. The bundle covariant Lagrange–Poincaré approach 871
 - 7.1. Covariant state space 871
 - 7.2. The variational principle 876
 - 7.3. The Kelvin–Noether Theorem 883
 - 7.4. Generalizations of the molecular strand 884
- 8. Outlook for further studies 885
- Appendix A. List of notations 887
- Appendix B. A modified Hamilton–Pontryagin approach 888
 - Appendix B.1. Filament dynamics 888
 - Appendix B.2. Nonlocal potential 890
- Appendix C. Formulation of nonlocal exact geometric rods in terms of quaternions 893
- Appendix D. The subgroup covariant Lagrange–Poincaré approach 898

1. Introduction

1.1. Physical setup

Long molecules are often modeled as strands of many individual charged units. Generally, the dynamics of such a charged molecular strands depends on both its local elastic deformations and the nonlocal (screened electrostatic) interactions of charged units across the loops in the strand. These electrostatic interactions depend on the spatial distances and relative orientations between the individual charged units at different locations along the strand.

The direct computational approach to such a complex problem is a full molecular dynamics simulation, using methods that take into account all (or most of) the forces between the atoms of the biological molecule as well as surrounding water molecules; see, for example, [33].

In contrast, many previous studies have addressed the elastic continuum dynamics of the charged strands using Kirchhoff’s approach [25]. For historical reviews and citations of this approach see, for example, [7,8]. A comprehensive survey of both the history and present state of the field can be found in [1]. Recent advances using this approach, especially in the context of helical structures, appear in, for

example, [3,14–17,34]. Although many important results have been obtained by the traditional continuum theory approach, it has a limitation. Namely, the generalization of the classical Kirchhoff theory to account for the torque caused by the long-range electrostatic interaction of molecules in different spatial locations along a flexible strand remains elusive, although the force due to electrostatic interaction has been captured by the traditional theory. See, for example, the article [7] which reviews progress in dynamical investigations of charged units distributed along a strand. In general, the lack of a consistent continuum model incorporating both torques and forces from electrostatic interactions has hampered analytical considerations; see for example [3] for additional discussion.

This paper introduces a framework that allows treatment of both torques and forces arising from electrostatic interactions. We should note that even in the absence of a continuum dynamical model for such nonlocal interactions, it is still possible to obtain *static* solutions using energy minimization techniques. For example, interesting helical static solutions of pressed elastic tubes using interactions that prevent self-intersection of the tubes were obtained in [4]. The difficulty in computing the dynamical effects of torque due to long-range interactions among the molecular subunits arises because the classical Kirchhoff theory is formulated in a frame moving with the strand, but it deals with a mixture of variables, some measured in the fixed spatial frame and some in the body frame. The torque due to long-range interactions presents a particular difficulty for the mixed representations in the Kirchhoff theory, because it is applied at base points of a curve that is moving in space. That is, the spatial Euclidean distances and relative orientations of the molecules must be reconstructed at each time step during the sinuous motion and twisting of the strand before any self-consistent computation can be made of the forces and torques due to long-range electrostatic interactions.

In fact, even when electrostatic forces are not involved, the motion of realistic curves in space is inherently nonlocal, because of the requirement that the curve not cross itself during the dynamics. In the purely elastic Kirchhoff approach, such nonlocal considerations are neglected. Physically, however, self-intersections are prevented by the existence of a short-range potential (for example, Lennard–Jones potential) that produces highly repulsive forces when two points along the curve approach each other. Thus, forces between segments of the strand that could be quite distant along its arc length are essential for the physical description of its dynamics.

This paper casts the problem of strand dynamics for an arbitrary intermolecular potential into the *convective representation* of continuum dynamics introduced in [22] and applied in exact geometric rod theory in [37]. The spatial and convective representations of continuum dynamics are the analogs, respectively, of the spatial and body representations of rigid body dynamics on $SO(3)$. This analogy arises because the configuration spaces for continuum dynamics with microstructure and for rigid bodies are both Lie groups. In both cases, the spatial velocities are right-invariant vector fields, while the convective, or body, velocities are the corresponding left-invariant vector fields.

If the curve were rigidly fixed in space, and the attached molecules on this fixed curve were simply allowed to rotate freely at each position, the theory of motion

based on nonlocal interaction between different molecules would be more straightforward. Of particular interest here is the work [31], where a single charge was attached at each point along a fixed filament by a rigid rod of constant length that was allowed to rotate in a transverse plane. These charges were allowed to interact locally with other nearby charges that were similarly attached to planar rotors of constant length mounted transversely to the fixed filament.

The model in [31] comprised a fixed base strand and rigid charge configurations described by $SO(2)$ (that is, one rotor in each normal plane). This model will be generalized here to allow flexible motion of the base strand (time-dependent bend, twist, writhe, and extension) while also including all the degrees of freedom of molecular orientation in $SO(3)$ excited during the process of, say, folding. According to this more general class of models, a long molecule is represented as a flexible filament or strand, along which are attached various different types of rigid conformations of sub-molecules that may *swivel* relative to each other in three dimensions under their mutual interactions. The flexibility of the filament arises physically because the electrostatic interaction between any pair of these rigid conformations, either along the filament or across from one loop to another of its folds, is much weaker than the internal interactions that maintain the shape of an individual charged conformation. The application of the present model to DNA is limited, however, because it fails to model the primary feature of DNA—its unzipping [32]. The further challenge of modeling unzipping of a double strand in the convective representation will be deferred to future work.

The primary aim of this paper is to formulate the dynamics of nonlocal interactions on a continuum strand carrying charged microstructure by using Lie symmetry reduction in the convective representation. This new formulation of symmetry-reduced nonlocal convective strand dynamics raises many interesting and nontrivial issues for future research. Among these issues are the classification and stability analysis of equilibrium solutions, dynamics of conformational changes (folding/unfolding) and formulation of computational approaches in the convective representation, all of which provide challenges for future research.

Scope of the paper. This paper considers rigid charge conformations (RCCs) that are mounted along a flexible moving filament. These RCCs are more complex than the planar pendula considered for a fixed base strand in [31]. They are mounted in orthonormal frames defined at each point along the strand. They are allowed to interact with each other via a nonlocal (for example, screened electrostatic, or Lennard–Jones) potential. Our model for the motion of the filament derives from the *geometrically exact rod theory* of Simó et al. [37], which is expressed in the *convective representation* of continuum mechanics. The rotations of rigid charge conformations along the flexible filament are illustrated in Fig. 1.

These rigid conformations of multiple charges interact via a nonlocal effective many-body potential representing their screened electrostatic interactions. The nonlocal interactions among these RCCs depend on their separations in the ambient space and their relative orientations, which are both allowed to evolve with the filament motion. Thus, the inertial motion of a pair of RCCs mounted at any two spatial points $\mathbf{r}(s, t)$ and $\mathbf{r}(s', t)$ along the filament is governed by an effective potential interaction energy that depends on their spatial separation and relative orientation.

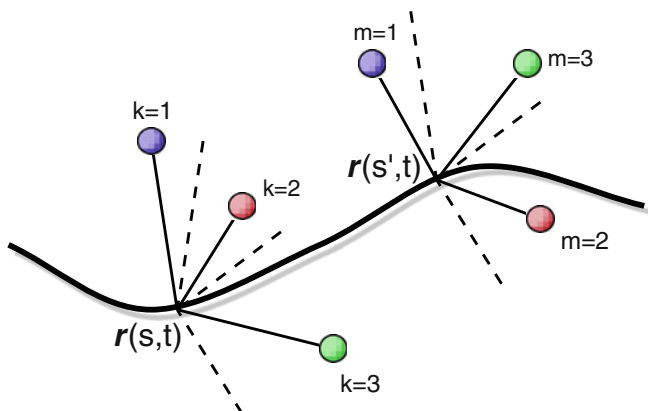


Fig. 1. Rigid conformations of charges are distributed along a curve. Note that this is a spatial representation of the orientations of these conformations of charges

The filament is taken to be one-dimensional, although the orientations of the rigid charged conformations mounted along it are three-dimensional. A practical example to which our filament approach would potentially apply is the vinylidene fluoride (VDF) oligomer [36], which may be approximated by a strand carrying a dipole moment whose orientation is perpendicular to the axis of the strand. The VDF oligomer strand is approximately straight for small lengths, but it forms complex shapes due to electrostatic interactions for longer lengths. In our framework, the undisturbed configuration of VDF polymer will correspond to a straight elastic filament, along which a rigid conformation of two opposite charges is positioned so that the dipole moment vector formed by those charges is perpendicular to the axis of the filament. The present paper is limited to formulating the geometrically exact model of the charged molecular strand and studying its mathematical structure. Many challenges remain to be investigated in future research concerning the properties and solution behavior of this model.

The dynamical influence of non-local electrostatic forces on rod mechanics is studied here using various approaches, including the Euler–Poincaré variational method [21]. This variational approach leads, for example, to an equivalent Lie–Poisson Hamiltonian formulation of the new equations appearing below in (3.23), (3.26). Applying the $\text{Ad}^*_{(\Lambda, r)^{-1}}$ transformation from convective to spatial variables in these equations streamlines their form and exposes the meaning of the interplay among their various local and nonlocal terms, relative to the Kirchhoff theory.

The convective formulation presented here applies equally well when the underlying substrate manifold (the filament, here) becomes multidimensional; so this formulation would also be applicable to such problems as the motion of charged sheets, or charged elastically deformable media. Although we present part of the relevant geometry here, we leave its applications in higher dimensions for a later publication.

Plan. The paper is written in two relatively independent, complementary parts that are meant to act as a “Rosetta stone” for expressing applications of symme-

try reduction for molecular strand dynamics from the following two perspectives. The first part of the paper consists of Sections 2–3, in which the reduced dynamics of the charged strands are derived by means of “bare hands” methods that use only variational principles and vector calculus. In contrast, the second part of the paper, consisting of Sections 4–7, contains a differential geometric perspective meant to elucidate the mathematical structure of the equations of motion derived in the first part. Additional information that enhances the formulation, but is not directly along the line of development of the rest of the paper, appears in the Appendices.

With this two-part organization of the paper, researchers in Molecular Dynamics, for example, who may be looking for a continuum model for DNA dynamics, can examine the equations of motion in the first part of the paper without being concerned about the differential geometry. Likewise, mathematicians can find the differential geometric structure in this model of strand dynamics without concerning themselves about how the model would apply in DNA experiments.

The paper is organized as follows. Section 1.2 outlines the content of the paper in mathematical terms by giving an overview of the various spatial representations of filament dynamics discussed here from the canonical and covariant point of views. Section 1.3 connects our results to the earlier literature. Paragraph 1.3.1 relates the theory presented here to the classical elastic rod approach pioneered by Kirchhoff. The need to keep track of spatial separations in long-range electrostatic interactions requires that we write the dynamics in either the spatial or convective representations, as opposed to using the Kirchhoff mixed representation. Paragraph 1.3.2 considers the simplified case when the orientations of the RCCs along the curve may depend on time, but the position of any point s along the curve is fixed, thereby connecting to earlier work in [31].

Section 2 incorporates the flexible motion of the filament into the dynamics by using the geometrically exact rod theory given in [37]. The equations of motion are derived in convective form in Section 3.1 by using the Euler–Poincaré approach, modified to allow for nonlocal interactions. A second derivation is given using the different, but equivalent, Hamilton–Pontryagin approach in Appendix B. The strand equations in the convective representation are formulated as conservation laws along the filament in Section 3.2 and their affine Lie–Poisson Hamiltonian structure is elucidated in Section 3.3. Appendix C presents the convective frame dynamics of the flexible strand in terms of quaternions, which we hope will be useful in numerical computations.

Section 4 introduces affine Lie group actions. Section 5 explains the background for the affine Euler–Poincaré and affine Lie–Poisson approaches and applies this framework to the dynamics of charged strands. Section 6 introduces a change of coordinates that decouples the equations into their horizontal and vertical parts in a principal bundle framework. Section 7 explains the geometric structure of this coordinate change and leads to the covariant Lagrange–Poincaré formulation. Section 7.4 and Appendix D discuss generalizations of the molecular strand to higher dimensions. Also in Appendix D, the equations of motion are obtained by an alternative covariant Lagrange–Poincaré approach. Section 8 summarizes our conclusions and sets out promising directions for further studies.

1.2. Mathematical setup

1.2.1. Description of the variables involved In the *Lagrangian representation*, the motion of a strand is described by the variables $\Lambda(s, t) \in SO(3)$ and $\mathbf{r}(s, t) \in \mathbb{R}^3$. The vector $\mathbf{r}(s, t)$ is the *spatial position of the filament* and the variable $\Lambda(s, t)$ denotes the *rotation of the RCC at the point s along the filament at time t* . Here $s \in [0, L]$ is a parameter spanning a fixed interval. The time and space derivatives yield, respectively, the *material velocity* ($\dot{\Lambda}(s, t), \dot{\mathbf{r}}(s, t)$) and the *angular and linear deformation gradients* ($\Lambda'(s, t), \mathbf{r}'(s, t)$). Given Λ and \mathbf{r} , we define notation for the following *reduced variables*

$$\begin{aligned}\Omega &= \Lambda^{-1} \Lambda' \in \mathfrak{so}(3), \\ \omega &= \Lambda^{-1} \dot{\Lambda} \in \mathfrak{so}(3), \\ \mathbf{\Gamma} &= \Lambda^{-1} \mathbf{r}' \in \mathbb{R}^3, \\ \boldsymbol{\gamma} &= \Lambda^{-1} \dot{\mathbf{r}} \in \mathbb{R}^3, \\ \boldsymbol{\rho} &= \Lambda^{-1} \mathbf{r} \in \mathbb{R}^3.\end{aligned}\tag{1.1}$$

Remark 1.1. (Notation) Quantities defined using derivatives in s are denoted using capital Greek letters, whereas lower-case Greek letters (except for $\boldsymbol{\rho}$) denote quantities whose definitions involve derivatives with respect to time. Bold letters, for example $\mathbf{\Gamma}$, denote vectors in \mathbb{R}^3 whereas Ω is a 3×3 skew-symmetric matrix in the Lie algebra $\mathfrak{so}(3)$.

Definition 1.1. The *hat map* $\hat{\cdot} : (\mathbb{R}^3, \times) \rightarrow (\mathfrak{so}(3), [\cdot, \cdot])$ is the Lie algebra isomorphism given by $\hat{\mathbf{u}}\mathbf{v} = \mathbf{u} \times \mathbf{v}$ for all $\mathbf{v} \in \mathbb{R}^3$.

Thus, in an orthonormal basis of \mathbb{R}^3 and $\mathbf{u} \in \mathbb{R}^3$, the 3×3 antisymmetric matrix $\mathbf{u} := \hat{\mathbf{u}} \in \mathfrak{so}(3)$ has entries

$$u_{jk} = (\hat{\mathbf{u}})_{jk} = -\varepsilon_{jkl} u^l.\tag{1.2}$$

Here the symbol ε_{jkl} with $j, k, l \in \{1, 2, 3\}$ denotes the totally antisymmetric tensor density with $\varepsilon_{123} = +1$ that defines the cross product of vectors in \mathbb{R}^3 . In what follows, we shall abbreviate this notation by writing $\Omega := \hat{\boldsymbol{\Omega}}$ and $\omega := \hat{\boldsymbol{\omega}}$.

The physical interpretation of the variables (1.1) is as follows: The variable $\boldsymbol{\rho}(s, t)$ represents the *position of the filament in space as viewed by an observer who rotates with the RCC at (s, t)* . The variables $(\boldsymbol{\Omega}(s, t), \mathbf{\Gamma}(s, t))$ describe the *deformation gradients as viewed by an observer who rotates with the RCC*. The variables $(\boldsymbol{\omega}(s, t), \boldsymbol{\gamma}(s, t))$ describe the *body angular velocity* and the *linear velocity as viewed by an observer who rotates with the RCC*.

1.2.2. The canonical point of view The canonical viewpoint of continuum dynamics derives the equations of motion by applying a process of reduction by symmetry to a phase space which is usually a cotangent bundle T^*Q endowed with a canonical symplectic form. Thus, at unreduced level, the motion is given by the *canonical Hamilton equations*

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q},$$

for a Hamiltonian function $H : T^*Q \rightarrow \mathbb{R}$, invariant under the action of the symmetry group. On the Lagrangian side, the motion is governed by the *Euler–Lagrange equations*

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0,$$

that produce the equations of motion by Lagrangian reduction. Here, L denotes the Lagrangian of the system, defined on the tangent bundle TQ of the configuration manifold Q .

This approach has been extensively studied for fluids; see for example [29] for the Hamiltonian description and [21] for the Lagrangian side. In hydrodynamics, Q is the product of a Lie group G and a representation space V on which the group acts linearly as $G \times V \rightarrow V$. The dual space, V^* , is the space of linearly advected quantities such as the mass density or the magnetic field. The associated process of reduction by symmetry under the action of G is called *Lie–Poisson reduction for semidirect products* [21]. For such systems (in the left version), we have the relations

$$\begin{aligned} \xi(t) &= g(t)^{-1} \dot{g}(t), \\ a(t) &= g(t)^{-1} a_{\text{ref}}, \end{aligned} \tag{1.3}$$

where $g(t) \in G$ is the Lagrangian motion, $\xi(t)$ is the convective velocity, and $a(t) \in V^*$ is the evolution of the advected quantity for a given initial condition a_{ref} (reference configuration). Note that $a(t)$ is also a convective quantity.

For the molecular strand we have $g = (\Lambda, \mathbf{r})$ and $a = (\mathbf{\Omega}, \mathbf{\Gamma}, \boldsymbol{\rho})$. However, the relations (1.1) cannot be recovered from (1.3) because the variables $(\mathbf{\Omega}, \mathbf{\Gamma}, \boldsymbol{\rho})$ are not linearly advected. Thus, a generalization of (1.3) is needed, in which $g \in G$ acts on $a \in V^*$ by an *affine action*. Such a generalization is given by the process of *affine Euler–Poincaré reduction* developed in the context of complex fluids in [12]. This theory, which we recall in Section 5, produces the relations

$$\begin{aligned} \xi(t) &= g(t)^{-1} \dot{g}(t), \\ a(t) &= g(t)^{-1} a_{\text{ref}} + c(g(t)^{-1}), \end{aligned} \tag{1.4}$$

where the additional term c is a group one-cocycle.¹

If we take $a_{\text{ref}} = 0$, then the advected quantity evolves in time as

$$a(t) = c(g(t)^{-1}).$$

Remarkably, the evolution of $(\mathbf{\Omega}, \mathbf{\Gamma}, \boldsymbol{\rho})$ in (1.1) is precisely of this form for a well chosen cocycle. The variables $(\Lambda(s, t), \mathbf{r}(s, t))$ are interpreted as *time-dependent curves in the infinite dimensional Lie group*

$$G = \mathcal{F}([0, L], SE(3))$$

¹ That is, c satisfies the property $c(fg) = c(f) + fc(g)$, where f acts on $c(g)$ by a left representation, as discussed in Section 5.

of all smooth functions on $[0, L]$ taking values in $SE(3) \simeq SO(3) \ltimes \mathbb{R}^3$, the special Euclidean group, comprising the semidirect product action of three-dimensional rotations $SO(3)$ and translations \mathbb{R}^3 as in equation (2.9).

Remark 1.2. The variables

$$(\omega, \gamma) = (\Lambda, \mathbf{r})^{-1}(\dot{\Lambda}, \dot{\mathbf{r}}),$$

their associated momenta

$$(\pi, \mu) := \left(\frac{\delta l}{\delta \omega}, \frac{\delta l}{\delta \gamma} \right),$$

and the affine advected variables (Ω, Γ, ρ) are all *convective quantities*; see [30]. In this context, convective quantities may also be called *body quantities*, since they are defined in a frame following the motion of the molecular strand.

In contrast, the variables

$$(\pi^{(S)}, \mathbf{p}^{(S)}) := \text{Ad}_{(\Lambda, \mathbf{r})}^* \left(\frac{\delta l}{\delta \omega}, \frac{\delta l}{\delta \gamma} \right) \quad (1.5)$$

are *spatial quantities*, that is, they are defined at fixed points in Euclidean space.

1.2.3. The covariant point of view The covariant point of view interprets the Lagrangian variables $(\Lambda(s, t), \mathbf{r}(s, t))$ as a map

$$(s, t) \mapsto (\Lambda(s, t), \mathbf{r}(s, t)), \quad (1.6)$$

in the group $SE(3)$ rather than as a curve

$$t \mapsto (\Lambda(\cdot, t), \mathbf{r}(\cdot, t)),$$

in the infinite dimensional configuration space $\mathcal{F}([0, L], SE(3))$. More precisely, the variables s and t are now treated in the same way and belong to the *spacetime* manifold $X := [0, L] \times \mathbb{R}$. Note that this is exactly the point of view taken in classical field theories. One should therefore see the map (1.6) as a *section* of the trivial fiber bundle

$$\pi_{XP} : P = X \times SE(3) \rightarrow X, \quad \pi_{XP}(x, \Lambda, \mathbf{r}) := (\Lambda, \mathbf{r}),$$

over the spacetime $X = [0, L] \times \mathbb{R} \ni (s, t) = x$. Indeed, by definition, a section σ of a bundle π_{XP} is a smooth map $\sigma : X \rightarrow P$ verifying the property $\pi_{XP} \circ \sigma = \text{id}_X$. Thus, since our bundle is trivial, a section reads

$$\sigma(x) = (x, \Lambda(x), \mathbf{r}(x)).$$

For the molecular strand, the Lagrangian depends on the map (1.6) as well as on its first partial derivatives. From the abstract covariant point of view this means that the Lagrangian \mathcal{L} depends on the first jet extension $j^1 \sigma(x) := T_x \sigma$ of the section

σ , where $T\sigma : TX \rightarrow TP$ denotes the tangent map of σ . More precisely, in our case we have

$$j^1\sigma(x) \cong (\Lambda(x), \mathbf{r}(x), \Lambda'(x)ds + \dot{\Lambda}(x)dt, \mathbf{r}'(x)ds + \dot{\mathbf{r}}(x)dt). \quad (1.7)$$

The first jet extension $j^1\sigma$ is in a natural way a section of a bundle over X called the *first jet bundle* $J^1P \rightarrow X$. Thus, the Lagrangian is abstractly a map

$$\mathcal{L} : J^1P \rightarrow \mathbb{R},$$

and the dynamics is given by the *covariant Euler–Lagrange equations*.

For the molecular strand, \mathcal{L} is $SO(3)$ -invariant, thus the first jet extension (1.7) yields the section

$$\left(\Lambda^{-1}\mathbf{r}, \Lambda^{-1}\Lambda' ds + \Lambda^{-1}\dot{\Lambda} dt, \Lambda^{-1}\mathbf{r}' ds + \Lambda^{-1}\dot{\mathbf{r}} dt \right),$$

by $SO(3)$ -reduction and one recovers the convective variables

$$(\boldsymbol{\rho}, \boldsymbol{\Omega} ds + \boldsymbol{\omega} dt, \boldsymbol{\Gamma} ds + \boldsymbol{\gamma} dt)$$

defined in (1.1). The precise geometric setting underlying this covariant reduction process will be explained in detail in Sections 6 and 7.

Reduction by the group $SO(3)$ yields a principal bundle structure on P given by

$$\pi_{\Sigma P} : P \rightarrow \Sigma := P/SO(3) = X \times \mathbb{R}^3, \quad \pi_{\Sigma P}(x, \Lambda, \mathbf{r}) = (x, \Lambda^{-1}\mathbf{r}).$$

This bundle over Σ should not be confused with the configuration bundle $\pi_{XP} : P \rightarrow X$ that has the same total space, but a different base. The following diagram illustrates the relationship among the different bundles arising in the covariant point of view:

$$\begin{array}{ccccc} \Lambda(s, t) \in SO(3) & & SO(3) & & \\ \downarrow & & \downarrow & & \\ (\Lambda, \mathbf{r})(s, t) \in SE(3) & \longrightarrow & P & \xrightarrow{\pi_{XP}} & X \\ \pi_{SE(3)} \downarrow & & \pi_{\Sigma P} \downarrow & & \downarrow \text{id} \\ \boldsymbol{\rho}(s, t) \in \mathbb{R}^3 & \longrightarrow & \Sigma & \xrightarrow{\pi_{X\Sigma}} & X. \end{array}$$

1.3. Connection to previous studies

1.3.1. Purely elastic motion and Kirchhoff equations for elastic rod The results of this paper in the convective representation may be compared to the classical Kirchhoff theory of the purely elastic rod, particularly in terms of the available conservation laws [1]. This comparison was presented in [37] for the purely elastic case, that is, the Lagrangian l is an explicit (local) function of the variables $l = l(\boldsymbol{\omega}, \boldsymbol{\gamma}, \boldsymbol{\Omega}, \boldsymbol{\Gamma}, \boldsymbol{\rho})$ defined in equation (1.1). The work of Simó et al. in [37] is extended here to the case of nonlocal interactions.

Of particular interest to us are the balance laws for angular and linear momenta. For this comparison, we shall use the notation of [7]. For simplicity, we assume that the position $\mathbf{r}(s)$ along the filament is given by the arc length s . This assumption conveniently avoids extra factors of $|\boldsymbol{\Gamma}(s)|$ in the expressions. We shall also mention here that in order to connect to the Kirchhoff theory, we need to make an explicit choice of $\Lambda(s) \in SO(3)$ as a transformation matrix from the fixed orthonormal basis $\{\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3\}$ of \mathbb{R}^3 to the orthonormal basis of directors $\{\mathbf{d}_1(s), \mathbf{d}_2(s), \mathbf{d}_3(s)\}$ describing the orientation of the filament (see Fig. 1). For this, we take

$$\mathbf{d}_i(s) = \Lambda_i^k(s) \mathbf{E}_k, \quad i = 1, 2, 3. \quad (1.8)$$

There is some ambiguity in the choice of the basis $\{\mathbf{d}_1(s), \mathbf{d}_2(s), \mathbf{d}_3(s)\}$ at every given point. The most popular selection of the basis is governed by the so-called *natural frame*. We shall not go into the details of this basis right now, and refer the reader to [7] for a more complete discussion. In principle, we need not have taken this particular choice of Λ , since for rigid charge conformations (RCC), the *relative* configuration of charges is not changed under the dynamics, and the configuration of an RCC state at each point s is completely described by a pair $(\Lambda, \mathbf{r}) \in SE(3)$. Taking Λ to be a different representation of an RCC would lead to a transformation $\Lambda(s, t) \mapsto A\Lambda(s, t)$ where $A \in SO(3)$ is a fixed matrix. While our description would be equivalent in this case, the explicit relation to Kirchhoff formulas would become less obvious.

We shall note that if the charge conformations were allowed to deform, then Λ would no longer be an element of $SO(3)$. Instead, the charge conformation would be described by a general invertible matrix Λ and a vector $\mathbf{r} \in \mathbb{R}^3$. No explicit relation to Kirchhoff's formulas would be possible in this case.

As mentioned in Section 1.1, Kirchhoff's approach precludes any simple computation of Euclidean distances between the charges, unless the spatial length-scale of the rigid charge conformations (RCCs) holding the charges at given point $\boldsymbol{\eta}_k(s)$ is negligible. It is interesting that in the more complex case considered here, the equations become formally equivalent to Kirchhoff's equations, provided the effects of non-locality are computed appropriately. In particular, one requires an appropriate mapping from the convective representation to the Kirchhoff representation, as well as some identities connecting nonlocal contributions to the *total* derivatives of the Lagrangian. This mapping is discussed in more detail in Section 3.2 below.

The linear momentum density \mathbf{p} is defined as $\mathbf{p}(s) = \rho_d(s) \dot{\mathbf{r}}(s)$, where $\rho_d(s)$ is the local mass density of the rod. In that case, the kinetic energy due to linear motion K_{lin} is given by

$$K_{\text{lin}} = \frac{1}{2} \int \rho_d(s) \|\dot{\mathbf{r}}(s)\|^2 ds = \frac{1}{2} \int \rho_d(s) \|\Lambda^{-1} \dot{\mathbf{r}}(s)\|^2 ds = \frac{1}{2} \int \rho_d(s) \|\boldsymbol{\gamma}(s)\|^2 ds.$$

Consequently, the variable \mathbf{p} and the linear momentum $\delta K_{\text{lin}} / \delta \boldsymbol{\gamma}$ are related by

$$\mathbf{p} = \rho_d \dot{\mathbf{r}} = \Lambda \rho_d \boldsymbol{\gamma} = \Lambda \frac{\delta K_{\text{lin}}}{\delta \boldsymbol{\gamma}}. \quad (1.9)$$

After these preliminaries, we are ready for a detailed comparison with Kirchhoff's theory. A point on a rod in Kirchhoff's theory is parameterized by

the distance $\mathbf{r}(s, t)$ measured from a *fixed* point in space. The i th component of the local angular momentum in the *body* frame $\{\mathbf{d}_1(s), \mathbf{d}_2(s), \mathbf{d}_3(s)\}$ is defined by $\pi^i(s) := \mathbb{I}_j^i(s)\omega^j(s)$, where $\omega^j(s)$ is the j th component of *body* angular velocity given by $\widehat{\omega}(s) := \omega(s) = \Lambda(s)^{-1}\dot{\Lambda}(s)$, and $\mathbb{I}_j^i(s)$ is the local value of the inertia tensor. Note that the inertia tensor $\mathbb{I}(s)$ expressed in body coordinates is time-independent. Thus the local kinetic energy due to rotation is given by

$$K_{\text{rot}} = \frac{1}{2} \int \omega(s) \cdot \mathbb{I}(s)\omega(s)ds.$$

Hence, the local body angular momentum density is given by

$$\pi = \mathbb{I}\omega = \frac{\delta K_{\text{rot}}}{\delta \omega}.$$

To write the conservation laws, we need to express the angular momentum in the fixed spatial frame $\{\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3\}$. To distinguish it from π which was expressed in the *body* frame $\{\mathbf{d}_1(s), \mathbf{d}_2(s), \mathbf{d}_3(s)\}$, we shall denote the same vector in the fixed spatial frame $\{\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3\}$ by $\pi^{(\mathbf{E})}$. This convention will be used for all other vectors. Thus, (1.8) yields

$$\pi(s) = \pi^i(s)\mathbf{d}_i(s) = \mathbb{I}_j^i(s)\omega^j(s)\mathbf{d}_i(s) = \mathbb{I}_j^i(s)\omega^j(s)\Lambda_i^k(s)\mathbf{E}_k = \pi^{(\mathbf{E}),k}(s)\mathbf{E}_k,$$

so the k th component of the spatial angular momentum is expressed in terms of the local body quantities $\mathbb{I}_j^i(s)$ and $\omega^k(s)$ as

$$\pi^{(\mathbf{E}),k} = \Lambda_i^k \mathbb{I}_j^i \omega^j = [\Lambda \mathbb{I} \omega]^k = \left[\Lambda \frac{\delta K_{\text{rot}}}{\delta \omega} \right]^k. \quad (1.10)$$

Therefore, the vector $\pi^{(\mathbf{E})}(s)$ of *body* angular momentum expressed in the *spatial* frame is connected to the local body quantities as

$$\pi^{(\mathbf{E})} = \Lambda \mathbb{I} \omega = \Lambda \frac{\delta K_{\text{rot}}}{\delta \omega}. \quad (1.11)$$

Remark 1.3. The vector $\pi^{(\mathbf{E})}$ and all other vectors with the superscript (\mathbf{E}) do not have the physical meaning of the angular momentum in the fixed frame. The true angular and linear momenta in the spatial frame will be denoted (see immediately below) with the superscript (S) . The quantities with the superscript (\mathbf{E}) are just the transformations of vectors with respect to rotation of the base frame. No confusion should arise over this distinction.

In general, it may be assumed for physical reasons that the Lagrangian in Kirchhoff's formulation has the form

$$l(\omega, \gamma, \Omega, \Gamma) = K_{\text{lin}}(\gamma) + K_{\text{rot}}(\omega) - E(\Omega, \Gamma), \quad (1.12)$$

where $E(\Omega, \Gamma)$ is a certain explicit function of Ω and Γ (not necessarily quadratic). In this case, the body forces $\mathbf{n} = \delta l / \delta \Gamma$ and torques $\mathbf{m} = \delta l / \delta \Omega$ are connected to the transformed quantities $\mathbf{n}^{(\mathbf{E})}$, $\mathbf{m}^{(\mathbf{E})}$ in Kirchhoff's theory as

$$\mathbf{n}^{(\mathbf{E})} = \Lambda \frac{\delta l}{\delta \Gamma}, \quad \mathbf{m}^{(\mathbf{E})} = \Lambda \frac{\delta l}{\delta \Omega}. \quad (1.13)$$

Next, we use formula (1.5) to transfer to spatial frame. Identifying elements of $\mathfrak{se}(3)^*$ with pairs of vectors $(\boldsymbol{\mu}, \boldsymbol{\eta}) \in \mathbb{R}^3$, produces a useful formula [30, 19] for the coadjoint action $SE(3) \times \mathfrak{se}(3)^* \rightarrow \mathfrak{se}(3)^*$, expressed in terms of vector cross products in \mathbb{R}^3 ,

$$\text{Ad}_{(\Lambda, \mathbf{r})}^* (\boldsymbol{\mu}, \boldsymbol{\eta}) = (\Lambda \boldsymbol{\mu} + \mathbf{r} \times \Lambda \boldsymbol{\eta}, \Lambda \boldsymbol{\eta}). \quad (1.14)$$

Thus, the spatial momenta—denoted by a superscript (S)—become

$$\begin{aligned} (\boldsymbol{\pi}^{(S)}, \mathbf{p}^{(S)}) &:= \text{Ad}_{(\Lambda, \mathbf{r})}^* \left(\frac{\delta l}{\delta \boldsymbol{\omega}}, \frac{\delta l}{\delta \boldsymbol{\gamma}} \right) = \left(\Lambda \frac{\delta l}{\delta \boldsymbol{\omega}} + \mathbf{r} \times \Lambda \frac{\delta l}{\delta \boldsymbol{\gamma}}, \Lambda \frac{\delta l}{\delta \boldsymbol{\gamma}} \right) \\ &= (\boldsymbol{\pi}^{(E)} + \mathbf{r} \times \mathbf{p}^{(E)}, \mathbf{p}^{(E)}), \end{aligned} \quad (1.15)$$

upon using (1.9) and (1.11). Analogously, using (1.13), the spatial torques $\mathbf{m}^{(S)}$ and forces $\mathbf{n}^{(S)}$ are expressed as

$$\begin{aligned} (\mathbf{m}^{(S)}, \mathbf{n}^{(S)}) &:= \text{Ad}_{(\Lambda, \mathbf{r})}^* \left(\frac{\delta l}{\delta \boldsymbol{\Omega}}, \frac{\delta l}{\delta \boldsymbol{\Gamma}} \right) = \left(\Lambda \frac{\delta l}{\delta \boldsymbol{\Omega}} + \mathbf{r} \times \Lambda \frac{\delta l}{\delta \boldsymbol{\Gamma}}, \Lambda \frac{\delta l}{\delta \boldsymbol{\Gamma}} \right) \\ &= (\mathbf{m}^{(E)} + \mathbf{r} \times \mathbf{n}^{(E)}, \mathbf{n}^{(E)}). \end{aligned} \quad (1.16)$$

The conservation laws in the Kirchhoff theory may now be written as

$$\frac{\partial}{\partial t} (\boldsymbol{\pi}^{(S)}, \mathbf{p}^{(S)}) + \frac{\partial}{\partial s} (\mathbf{m}^{(S)}, \mathbf{n}^{(S)}) = (\mathbf{T}, \mathbf{f}), \quad (1.17)$$

where \mathbf{T} and \mathbf{f} are external torques and forces, respectively. Equations (1.17) give, componentwise, the following linear and angular momentum conservation laws (cf. equations (2.5.5) and (2.5.7) of [7])

$$\frac{\partial}{\partial t} \mathbf{p}^{(E)} + \frac{\partial}{\partial s} (\mathbf{n}^{(E)} - \mathbf{F}) = 0, \quad (1.18)$$

$$\frac{\partial}{\partial t} (\boldsymbol{\pi}^{(E)} + \mathbf{r} \times \mathbf{p}^{(E)}) + \frac{\partial}{\partial s} (\mathbf{m}^{(E)} + \mathbf{r} \times \mathbf{n}^{(E)} - \mathbf{L}) = 0, \quad (1.19)$$

where \mathbf{F} and \mathbf{L} are defined as the indefinite integrals,

$$\mathbf{F} = \int^s \mathbf{f}(q) dq \quad \text{and} \quad \mathbf{L} = \int^s [\mathbf{r}(q) \times \mathbf{f}(q) + \mathbf{T}(q)] dq.$$

Opening the brackets in (1.18) and (1.19) gives the balances of linear and angular momenta in Kirchhoff's approach (cf. equations (2.3.5) and (2.3.6) of [7])

$$\frac{\partial \mathbf{p}^{(E)}}{\partial t} + \frac{\partial \mathbf{n}^{(E)}}{\partial s} = \mathbf{f}, \quad (1.20)$$

$$\frac{\partial \boldsymbol{\pi}^{(E)}}{\partial t} + \frac{\partial \mathbf{m}^{(E)}}{\partial s} + \frac{\partial \mathbf{r}}{\partial s} \times \mathbf{n}^{(E)} = \mathbf{T}. \quad (1.21)$$

To see how these Kirchhoff balance laws look in our convective representation, one may substitute relations (1.15) and (1.16) into (1.17) to obtain:

$$\frac{\partial}{\partial t} \left[\text{Ad}_{(\Lambda, \mathbf{r})}^* \left(\frac{\delta l}{\delta \boldsymbol{\omega}}, \frac{\delta l}{\delta \boldsymbol{\gamma}} \right) \right] + \frac{\partial}{\partial s} \left[\text{Ad}_{(\Lambda, \mathbf{r})}^* \left(\frac{\delta l}{\delta \boldsymbol{\Omega}}, \frac{\delta l}{\delta \boldsymbol{\Gamma}} \right) \right] = (\mathbf{T}, \mathbf{f}). \quad (1.22)$$

Assume now that the Lagrangian l depends explicitly on the additional variable $\rho = \Lambda^{-1}\mathbf{r}$. This dependence corresponds to *potential* forces exerting forces and torques. As shown in Section 3.2, in our representation the external torques \mathbf{T} and forces \mathbf{f} are given by

$$(\mathbf{T}, \mathbf{f}) = \text{Ad}_{(\Lambda, \mathbf{r})}^* \left(\frac{\delta l}{\delta \rho} \times \rho, \frac{\delta l}{\delta \rho} \right). \quad (1.23)$$

By using formula (1.14), relationship (1.23) expands, then simplifies to

$$\begin{aligned} \text{Ad}_{(\Lambda, \mathbf{r})}^* \left(\frac{\delta l}{\delta \rho} \times \rho, \frac{\delta l}{\delta \rho} \right) &= \left(\Lambda \left(\frac{\delta l}{\delta \rho} \times \rho \right) + \mathbf{r} \times \Lambda \frac{\delta l}{\delta \rho}, \Lambda \frac{\delta l}{\delta \rho} \right) \\ &= \left(\left(\Lambda \frac{\delta l}{\delta \rho} \right) \times (\Lambda \rho) + \mathbf{r} \times \Lambda \frac{\delta l}{\delta \rho}, \Lambda \frac{\delta l}{\delta \rho} \right) \\ &= \left(0, \Lambda \frac{\delta l}{\delta \rho} \right), \end{aligned} \quad (1.24)$$

upon recalling from (1.1) that $\Lambda \rho = \mathbf{r}$.

Remark 1.4. (Potential external forces produce no net torque) Equation (1.24) implies that potential external forces produce no net torque on the strand. Hence, the nonzero torques \mathbf{T} in (1.22) must arise from non-potential forces.

The conservation law (1.22) is formally equivalent to the classical expressions in (1.18) and (1.19), even if nonlocal interaction is present. This equivalence shows how the classical results (1.18) and (1.19) generalize for the case of nonlocal orientation-dependent interactions. Clearly, the conservation laws are simpler in the Kirchhoff representation. However, if nonlocal interactions are present (called *self-interaction forces* in [7]), the computation of the required time-dependent Euclidean distances in the interaction energy becomes problematic in the classical Kirchhoff approach. As we shall see below in Section 3.2, these conservation laws may be obtained, even when nonlocal interactions are present. Also in Section 3.2, we show that the nonlocal forces are included in the conservation law (1.22) and are expressed in the same form as a purely elastic conservation law.

The balance laws (1.18) and (1.19) are much simpler in appearance than the expressions in (1.22), as they do not involve computing (Λ, \mathbf{r}) at each instant in time and point in space. Thus for elastic rods, in the absence of nonlocal interactions, the Kirchhoff mixed (convective-spatial) representation appears simpler than in either the convective or spatial representations. However, the presence of nonlocal terms summons the more general convective approach introduced for this problem in [23].

Remark 1.5. (Reduction of static equations of motion to the heavy top). A famous analogy exists between the stationary shapes of an elastic filament and the equations of motion of a heavy top [24, 35]. This shows that the geometric approach also applies to the problem of determining the steady equilibrium solutions of filament dynamics. This paper focuses, however, on the derivations and geometric structures underlying the dynamical equations, rather than on the solutions of the equations.

1.3.2. Reductions for a fixed filament We begin by applying the ideas of the present paper to the particular case of a fixed filament, in order to compare the motion equations with those arising in [31]. This is achieved by a direct reduction of the more general case to the non-moving filament, and the rotation of the RCC by a general $SO(3)$ group element. Such a system is a generalization of the system presented in [31] from $SO(2)$ to $SO(3)$ rotations.

The analysis of filament dynamics driven by nonlocal interactions simplifies in the case when the position of the filament is fixed as $\mathbf{r}(s)$ and does not depend on time. For simplicity, we shall assume that the filament is straight and s is the arc length, so that $\mathbf{r}(s) = (s, 0, 0)^T$. The following reduced Lagrangian is invariant under the left action of the Lie group $SO(3)$:

$$l = \underbrace{\frac{1}{2} \int \omega(s) \cdot \mathbb{I}(s) \omega(s) ds}_{\text{Kinetic energy}} - \underbrace{\frac{1}{2} \int f(\boldsymbol{\Omega}(s)) ds}_{\text{Elastic energy}} - \underbrace{\frac{1}{2} \iint U(\boldsymbol{\rho}(s), \xi(s, s')) ds ds'}_{\text{Potential energy}}. \quad (1.25)$$

A nonlocal interaction term appears in the potential energy of relative orientation in this Lagrangian. This term involves a variable

$$\xi(s, s') = \Lambda^{-1}(s) \Lambda(s') \in SO(3),$$

which defines the relative orientation of rigid charge conformations at two *different points* in space. The variable $\xi(s, s') \in SO(3)$ is invariant with respect to simultaneous rotations of the coordinate frames for s and s' , but it is not an element of a Lie algebra. In particular, $\xi(s, s')$ is not a vector. The presence of nonlocal interactions introduces dependence on relative orientation, and thereby produces new types of nonlocal terms in the corresponding Euler–Poincaré dynamics obtained in applying reduction by $SO(3)$ symmetry to Hamilton’s principle.

Euler–Poincaré dynamics. Euler–Poincaré dynamics for the angular dynamics on a *fixed* filament follows from stationarity of the left invariant total action

$$S = \int l(\boldsymbol{\omega}, \boldsymbol{\rho}, \xi, \boldsymbol{\Omega}) dt.$$

Note that this case does not require computation of the evolution equation for $\boldsymbol{\gamma}$, since the filament is assumed to be fixed in space, that is, $\boldsymbol{\gamma} = \Lambda^{-1} \dot{\mathbf{r}} = 0$. The variational derivative δS for such a Lagrangian is computed as,

$$\delta S = \int \delta l(\boldsymbol{\omega}, \Lambda, \boldsymbol{\Omega}) dt = \int \left(\left\langle \frac{\delta l}{\delta \boldsymbol{\omega}}, \delta \boldsymbol{\omega} \right\rangle + \left\langle \Lambda^{-1} \frac{\delta l}{\delta \Lambda}, \Sigma \right\rangle + \left\langle \frac{\delta l}{\delta \boldsymbol{\Omega}}, \delta \boldsymbol{\Omega} \right\rangle \right) dt, \quad (1.26)$$

for the notation $\Sigma = \Lambda^{-1} \delta \Lambda$. As we will see in Section 3.1.1, these variations are related by

$$\begin{aligned} \delta \boldsymbol{\omega} &= \dot{\Sigma} + [\boldsymbol{\omega}, \Sigma] = \dot{\Sigma} + \text{ad}_{\boldsymbol{\omega}} \Sigma, \\ \delta \boldsymbol{\Omega} &= \Sigma' + [\boldsymbol{\Omega}, \Sigma] = \Sigma' + \text{ad}_{\boldsymbol{\Omega}} \Sigma, \\ \delta \boldsymbol{\rho} &= -\Sigma \times \boldsymbol{\rho}. \end{aligned}$$

Substituting these formulas into (1.26) then integrating by parts in the time t and one-dimensional coordinate s along the fiber yields

$$\begin{aligned} \delta S = \int \delta l \, dt = \int \left\langle -\frac{\partial}{\partial t} \frac{\delta l}{\delta \omega} + \text{ad}_\omega^* \frac{\delta l}{\delta \omega} - \frac{\partial}{\partial s} \frac{\delta l}{\delta \Omega} + \text{ad}_\Omega^* \frac{\delta l}{\delta \Omega} \right. \\ \left. - \int \left(-\frac{\partial U}{\partial \xi}(s, s') \xi^T(s, s') + \xi(s, s') \left(\frac{\partial U}{\partial \xi}(s, s') \right)^T \right) ds' \right. \\ \left. + \left(\frac{\delta l}{\delta \rho} \times \rho \right), \Sigma \right\rangle dt, \end{aligned} \quad (1.27)$$

where

$$\frac{\delta l}{\delta \rho} = -\frac{1}{2} \int \frac{\partial U}{\partial \rho}(\rho(s), \xi(s, s')) \, ds'.$$

Thus, Hamilton's principle $\delta S = 0$ implies the *Euler–Poincaré* equations,

$$\begin{aligned} -\frac{\partial}{\partial t} \frac{\delta l}{\delta \omega} + \text{ad}_\omega^* \frac{\delta l}{\delta \omega} = \frac{\partial}{\partial s} \frac{\delta l}{\delta \Omega} - \text{ad}_\Omega^* \frac{\delta l}{\delta \Omega} - \left(\frac{\delta l}{\delta \rho} \times \rho \right)^\wedge \\ + \int \left(-\frac{\partial U}{\partial \xi}(s, s') \xi^T(s, s') + \xi(s, s') \left(\frac{\partial U}{\partial \xi}(s, s') \right)^T \right) ds'. \end{aligned} \quad (1.28)$$

Note that these Euler–Poincaré equations are *nonlocal*. That is, they are integral-partial differential equations.

Reformulating (1.28) in terms of vectors yields the following generalization of equations considered by [31], written in a familiar vector form:

$$\begin{aligned} \left(-\frac{d}{dt} \frac{\delta l}{\delta \omega} + \frac{\delta l}{\delta \omega} \times \omega - \frac{\partial}{\partial s} \frac{\delta l}{\delta \Omega} - \Omega \times \frac{\delta l}{\delta \Omega} - \rho \times \frac{\delta l}{\delta \rho} \right)^\wedge \\ = \int \left(-\frac{\partial U}{\partial \xi}(s, s') \xi^T(s, s') + \xi(s, s') \left(\frac{\partial U}{\partial \xi}(s, s') \right)^T \right) ds'. \end{aligned} \quad (1.29)$$

In order to close the system, one computes the time derivative of $\xi(s, s') = \Lambda^{-1}(s')\Lambda(s)$:

$$\begin{aligned} \dot{\xi}(s, s') &= -\Lambda^{-1}(s')\dot{\Lambda}(s')\Lambda^{-1}(s')\Lambda(s) + \Lambda^{-1}(s')\dot{\Lambda}(s) \\ &= -\omega(s')\xi(s, s') + \xi(s, s')\omega(s). \end{aligned} \quad (1.30)$$

This expression is not quite a commutator because different positions s and s' appear in ω . However, operating with ξ^{-1} from the left in equation (1.30) gives a proper Lie-algebraic expression for the reconstruction of the relative orientation,

$$\xi^{-1}\dot{\xi}(s, s') = \omega(s) - \text{Ad}_{\xi^{-1}(s, s')}\omega(s'). \quad (1.31)$$

Formulas (1.28)–(1.30) generalize the results in [31] for a fixed filament from $SO(2)$ to $SO(3)$ rotations.

2. Motion of exact self-interacting geometric rods

2.1. Problem set-up

Suppose each rigid conformation of charges is identical and the k th electrical charge is positioned near a given spatial point \mathbf{r} through which the curve of base points of the RCCs passes. This curve is parametrized by a variable s which need not be the arc length. Rather, we take $s \in [0, L]$ to be a parameter spanning a fixed interval.²

The spatial reference (undisturbed) state for the k th charge in a given RCC is the sum $\mathbf{r}(s) + \boldsymbol{\eta}_k(s)$. That is, $\boldsymbol{\eta}_k(s)$ is a vector of constant length that determines the position of the k th electrical charge relative to the point $\mathbf{r}(s)$ along the curve in its reference configuration. The $\boldsymbol{\eta}_k(s)$ specify the shape of the rigid conformation of charges. At time t the position \mathbf{c}_k of the k th charge in the rigid conformation anchored at spatial position $\mathbf{r}(s, t)$ along the curve parametrized by s may rotate to a new position corresponding to the orientation $\Lambda(s, t)$ in the expression

$$\mathbf{c}_k(s, t) = \mathbf{r}(s, t) + \Lambda(s, t)\boldsymbol{\eta}_k(s), \quad \text{where } \Lambda(s, 0) = \text{Id}. \quad (2.1)$$

This rigid conformational rotation is illustrated in Fig. 1. In Mezic's case [31], the rotation is in the plane, so that $\Lambda \in SO(2)$, and there is only one charge, so $k = 1$. From now on we shall suppress notation for time dependence without danger of confusion.

2.2. Convective representation of nonlocal potential energy

One part of the potential energy of interaction between rigid conformations of charges at spatial coordinates $\mathbf{r}(s)$ and $\mathbf{r}(s')$ along the filament depends only on the magnitude $|\mathbf{c}_m(s') - \mathbf{c}_k(s)|$ of the vector from charge k at spatial position $\mathbf{c}_k(s)$ to charge m at spatial position $\mathbf{c}_m(s')$. This is the Euclidean spatial distance

$$d_{k,m}(s, s') = |\mathbf{c}_m(s') - \mathbf{c}_k(s)| \quad (2.2)$$

between the k th and m th charges in the two conformations whose base points are at $\mathbf{r}(s)$ and $\mathbf{r}(s')$, respectively. In this notation, the potential energy is given by

$$E = E_{\text{loc}}(\boldsymbol{\Omega}, \boldsymbol{\Gamma}) - \sum_{k,m} \frac{1}{2} \iint U(d_{k,m}(s, s')) \left| \frac{d\mathbf{r}}{ds}(s) \right| \left| \frac{d\mathbf{r}}{ds}(s') \right| ds ds', \quad (2.3)$$

for an appropriate physical choice of the *interparticle interaction potential* $U(d_{k,m})$, and the quantities $\boldsymbol{\Omega}, \boldsymbol{\Gamma}$ (and $\boldsymbol{\omega}, \boldsymbol{\gamma}, \boldsymbol{\rho}$ below) are defined in (1.1). The part $E_{\text{loc}}(\boldsymbol{\Omega}, \boldsymbol{\Gamma})$ represents the purely elastic part of the potential and is usually taken to be a quadratic function of the deformations $(\boldsymbol{\Omega}, \boldsymbol{\Gamma})$, but more complex expressions are possible

² Note: limiting its parametrization to a fixed interval does not mean that the filament is inextensible.

as well; we shall not restrict the functional form of that dependence. The total Lagrangian l is then written as the sum of a local l_{loc} and a nonlocal l_{np} part:

$$l_{\text{loc}} = K(\boldsymbol{\omega}, \boldsymbol{\gamma}) - E_{\text{loc}}(\boldsymbol{\Omega}, \boldsymbol{\Gamma}, \boldsymbol{\rho}) \quad \text{and} \quad l_{\text{np}} = -E_{\text{np}}, \quad (2.4)$$

where K is the kinetic energy that depends only on the local velocities $\boldsymbol{\omega}, \boldsymbol{\gamma}$. For the sake of generality, here and everywhere else below, we shall simply consider the total Lagrangian to be a sum of the local part $l_{\text{loc}}(\boldsymbol{\omega}, \boldsymbol{\gamma}, \boldsymbol{\Omega}, \boldsymbol{\Gamma}, \boldsymbol{\rho})$, and the nonlocal part given by (2.3):

$$l = l_{\text{loc}}(\boldsymbol{\omega}, \boldsymbol{\gamma}, \boldsymbol{\Omega}, \boldsymbol{\Gamma}, \boldsymbol{\rho}) + l_{\text{np}}. \quad (2.5)$$

The scalar distance $d_{k,m}$ in (2.2) and (2.3) may also be expressed in terms of vectors seen from the frame of orientation of the rigid body at a spatial point $\mathbf{r}(s)$ along the filament, as

$$\begin{aligned} d_{k,m}(s, s') &= |\mathbf{c}_m(s') - \mathbf{c}_k(s)| \\ &= \left| \Lambda^{-1}(s) (\mathbf{c}_m(s') - \mathbf{c}_k(s)) \right| \\ &= \left| \Lambda^{-1}(s) \left(\mathbf{r}(s') - \mathbf{r}(s) + \Lambda^{-1}(s) \Lambda(s') \boldsymbol{\eta}_m(s') - \boldsymbol{\eta}_k(s) \right) \right| \\ &=: |\boldsymbol{\kappa}(s, s') + \xi(s, s') \boldsymbol{\eta}_m(s') - \boldsymbol{\eta}_k(s)|, \end{aligned} \quad (2.6)$$

where we have defined the quantities

$$\boldsymbol{\kappa}(s, s') := \Lambda^{-1}(s) (\mathbf{r}(s') - \mathbf{r}(s)) \in \mathbb{R}^3 \quad \text{and} \quad \xi(s, s') := \Lambda^{-1}(s) \Lambda(s') \in SO(3). \quad (2.7)$$

The first of these quantities is the spatial vector from $\mathbf{r}(s)$ to $\mathbf{r}(s')$, as seen from the orientation $\Lambda(s)$ of the rigid charge conformation located at coordinate label s along the filament. The second is the relative orientation of the rigid charge conformations located at coordinate labels s and s' . For later use, we record the *transposition identities*,

$$\xi(s, s')^T = \xi(s', s) = \xi(s, s')^{-1}, \quad (2.8)$$

which follow from the definition of $\xi(s, s')$ in (2.7).

Remark 2.1. (Left $SO(3)$ invariance) Both the body separation vector $\boldsymbol{\kappa}(s, s')$ and the relative orientation $\xi(s, s')$ defined in (2.7) are invariant under rotations of the spatial coordinate system obtained by the left action

$$(\mathbf{r}(s') - \mathbf{r}(s)) \mapsto O (\mathbf{r}(s') - \mathbf{r}(s)) \quad \text{and} \quad \Lambda \mapsto O \Lambda,$$

by any element O of the proper rotation group $SO(3)$.

Proposition 2.1. (Left $SE(3)$ invariance) *The quantities $(\xi, \boldsymbol{\kappa}) \in SO(3) \times \mathbb{R}^3$ defined in (2.7) are invariant under all transformations of the special Euclidean group $SE(3)$ acting on the left.*

Proof. As a set, the special Euclidean group $SE(3)$ is the Cartesian product $SE(3) = SO(3) \times \mathbb{R}^3$ whose elements are denoted as (Λ, \mathbf{r}) . Its group multiplication is given, for example, in [19] by the *semidirect product action*,

$$(\Lambda_1, \mathbf{r}_1)(\Lambda_2, \mathbf{r}_2) = (\Lambda_1 \Lambda_2, \mathbf{r}_1 + \Lambda_1 \mathbf{r}_2), \quad (2.9)$$

where the action of $\Lambda \in SO(3)$ on $\mathbf{r} \in \mathbb{R}^3$ is denoted as the concatenation $\Lambda \mathbf{r}$ and the other notation is standard. For the choice

$$(\Lambda_1, \mathbf{r}_1) = (\Lambda, \mathbf{r})^{-1}(s) \quad \text{and} \quad (\Lambda_2, \mathbf{r}_2) = (\Lambda, \mathbf{r})(s'),$$

the $SE(3)$ multiplication rule (2.9) yields the quantities $(\xi, \kappa) \in SO(3) \times \mathbb{R}^3$ as

$$(\Lambda, \mathbf{r})^{-1}(s)(\Lambda, \mathbf{r})(s') = (\xi(s, s'), \kappa(s, s')). \quad (2.10)$$

This expression is invariant under the left action $(\Lambda, \mathbf{r}) \mapsto (O, \mathbf{v})(\Lambda, \mathbf{r})$ of any element (O, \mathbf{v}) of the special Euclidean group $SE(3)$. \square

Remark 2.2. The $SE(3)$ setting will be especially important to the development of the Lagrange–Poincaré formulation of the dynamical filament equations in Section 7.

Next, let us define the following $SE(3)$ -invariant quantities, where prime denotes the derivative with respect to s and dot is the derivative with respect to t :

$$\begin{aligned} \Omega &= \Lambda^{-1} \Lambda' && \in \mathfrak{so}(3), \\ \omega &= \Lambda^{-1} \dot{\Lambda} && \in \mathfrak{so}(3), \\ \Gamma &= \Lambda^{-1} \mathbf{r}' && \in \mathbb{R}^3, \\ \gamma &= \Lambda^{-1} \dot{\mathbf{r}} && \in \mathbb{R}^3, \\ \rho &= \Lambda^{-1} (\mathbf{r} - \mathbf{r}_0(s)) && \in \mathbb{R}^3. \end{aligned} \quad (2.11)$$

Hereafter, we shall choose $\mathbf{r}_0(s) = \mathbf{0}$ to recover the bundle coordinates (1.1).

Remark 2.3. Note that here $\Lambda, \mathbf{r}, \Omega, \omega, \Gamma, \gamma, \rho$ are interpreted as functions of the two variables s and t . It will be important to see these variables as time-dependent curves with values in function spaces. For example, we can interpret $\Lambda(s, t)$ as a function of space and time

$$(s, t) \in [0, L] \times \mathbb{R} \mapsto \Lambda(s, t) \in SO(3),$$

or we can see Λ as a curve

$$t \in \mathbb{R} \mapsto \Lambda(\cdot, t) \in \mathcal{F}([0, L], SO(3)),$$

in the group $\mathcal{F}([0, L], SO(3))$ of all smooth functions defined on $[0, L]$ with values in $SO(3)$.

This observation is fundamental and leads to two different geometric approaches to the same equations: the *affine Euler–Poincaré* and the *covariant Lagrange–Poincaré* approaches.

Note that the group operation $\mathcal{F}([0, L], SO(3))$ is given by pointwise multiplication, and that the space $\mathcal{F}([0, L], SO(3))$ can be endowed with the structure of an infinite dimensional Fréchet Lie group. We refer to [26] for an account of Fréchet Lie groups in the framework of manifolds of maps from the point of view of the convenient calculus.

Remark 2.4. Since $\Lambda \in SO(3)$, one finds that

$$\left| \frac{d\mathbf{r}}{ds}(s) \right| = \left| \Lambda^{-1} \frac{d\mathbf{r}}{ds}(s) \right| = |\boldsymbol{\Gamma}(s)|, \quad (2.12)$$

and the nonlocal part of the potential energy in (2.3) reduces to

$$E_{\text{np}} = - \sum_{k,m} \frac{1}{2} \iint U(d_{k,m}(s, s')) |\boldsymbol{\Gamma}(s)| |\boldsymbol{\Gamma}(s')| ds ds'. \quad (2.13)$$

Remark 2.5. Everywhere in this paper, we shall assume that the nonlocal part of the Lagrangian l_{np} is a function or functional of $\boldsymbol{\Gamma}$, ξ , and $\boldsymbol{\kappa}$. It could, for example, be expressed in the integral form

$$l_{\text{np}}(\xi, \boldsymbol{\kappa}, \boldsymbol{\Gamma}) = \iint U(\xi(s, s'), \boldsymbol{\kappa}(s, s'), \boldsymbol{\Gamma}(s), \boldsymbol{\Gamma}(s')) ds ds' \quad (2.14)$$

or be a more general functional. In this work, we shall consistently use formula (2.14) to make our computations more explicit although, of course, the methods would apply to more general functionals. Clearly, expression (2.3) is a reduction of (2.14) obtained when the energy of the system of charges is a (half-)sum of interactions between all charges. This happens, for example, when investigating electrostatic or screened electrostatic charges in a linear medium.

Even though the expression $l_{\text{np}} = l_{\text{np}}(\xi, \boldsymbol{\kappa}, \boldsymbol{\Gamma})$ is rather general, it is interesting to note that physical systems exist whose nonlocal interactions are more complex. For example, the electrostatic potential around a DNA molecule immersed in a fluid satisfies the nonlinear Poisson–Boltzmann equation and finding the potential in that case is a well-known problem for supercomputers [2]. Direct analytical solution of this equation is usually impossible except for some idealized cases. However, even in this case, one can consider the nonlocal Lagrangian as a functional of the variables $(\xi, \boldsymbol{\kappa}, \boldsymbol{\Gamma})$, and compute (at least in principle) the variational derivatives required by our theory. In that case, our theory is applicable as well.

2.3. Kinematics

We first define auxiliary kinematic equations that hold without any reference to dynamics. We call these *advection relations*, in order to distinguish them from the *dynamical equations* (derived later) that balance the forces determined from the physics of the problem. In contrast, the advection relations hold for all strands, irrespective of their dynamic properties.

In order to derive the first set of advection relations, we compute the time and space derivatives of $\boldsymbol{\rho}(s, t) = \Lambda(s, t)^{-1} \mathbf{r}(s, t)$. First, the s -derivative along the filament is given by:

$$\boldsymbol{\rho}' = -\Lambda^{-1} \Lambda' \Lambda^{-1} \mathbf{r} + \Lambda^{-1} \mathbf{r}',$$

and hence equations (2.11) imply

$$\boldsymbol{\rho}' = -\boldsymbol{\Omega} \boldsymbol{\rho} + \boldsymbol{\Gamma} = -\boldsymbol{\Omega} \times \boldsymbol{\rho} + \boldsymbol{\Gamma}. \quad (2.15)$$

Next, the time derivative is written as

$$\dot{\boldsymbol{\rho}} = -\Lambda^{-1} \dot{\Lambda} \Lambda^{-1} \mathbf{r} + \Lambda^{-1} \dot{\mathbf{r}} \quad (2.16)$$

and equations (2.11) yield the formula,

$$\dot{\boldsymbol{\rho}} = -\boldsymbol{\omega} \boldsymbol{\rho} + \boldsymbol{\gamma} = -\boldsymbol{\omega} \times \boldsymbol{\rho} + \boldsymbol{\gamma}. \quad (2.17)$$

The next set of advection relations is derived by the equality of cross-derivatives with respect to t and s for any sufficiently smooth quantity. First, we use the fact that $\partial_s \partial_t \mathbf{r} = \partial_t \partial_s \mathbf{r}$. Equality of these cross-derivatives implies the relations

$$\boldsymbol{\gamma}' = -\boldsymbol{\Omega} \times \boldsymbol{\gamma} + \Lambda^{-1} \dot{\mathbf{r}}'$$

and

$$\dot{\boldsymbol{\Gamma}} = -\boldsymbol{\omega} \times \boldsymbol{\Gamma} + \Lambda^{-1} \dot{\mathbf{r}}'.$$

The difference of the last two equations yields the following relation

$$\dot{\boldsymbol{\Gamma}} + \boldsymbol{\omega} \times \boldsymbol{\Gamma} = \boldsymbol{\gamma}' + \boldsymbol{\Omega} \times \boldsymbol{\gamma}. \quad (2.18)$$

As we shall see later, the latter is a type of *zero-curvature relation*. Similarly, equality of the cross-derivatives $\partial_s \partial_t \Lambda = \partial_t \partial_s \Lambda$ yields the other advection relation:

$$\dot{\boldsymbol{\Omega}} - \boldsymbol{\omega}' = \boldsymbol{\omega} \times \boldsymbol{\Omega}. \quad (2.19)$$

The differential geometric meaning of these zero-curvature relations will be discussed further in Section 7.2.6.

2.4. Remark on n -dimensional generalization and use of other groups

The previous setting may be generalized to n dimensions and to arbitrary Lie groups. This is useful not only for the generalization of charged strands to membranes and, more generally, to deformable media; it also gives a more transparent vision of the underlying geometric structure underlying the phenomena.

Consider the semidirect product $\mathcal{O} \ltimes E$ of a Lie group \mathcal{O} (whose Lie algebra is denoted by \mathfrak{o}) with a *left* representation space E . The variables \mathbf{r} and Λ defined above are now functions defined on a spacetime $\mathcal{D} \times \mathbb{R}$, where \mathcal{D} is a n -dimensional manifold:

$$\Lambda : (s, t) \in \mathcal{D} \times \mathbb{R} \mapsto \Lambda(s, t) \in \mathcal{O}, \quad \text{and} \quad r : (s, t) \in \mathcal{D} \times \mathbb{R} \mapsto r(s, t) \in E.$$

We will avoid using boldface notation as the functions we consider may be more general geometric quantities, not only vectors. As before, “dot” (·) over a quantity denotes its time derivative. The derivative with respect to a variable in \mathcal{D} is denoted by \mathbf{d} ; for $\mathcal{D} = [0, L]$ this was previously denoted by “prime” (′). The definitions (2.11) become

$$\begin{aligned}\Omega &= \Lambda^{-1} \mathbf{d}\Lambda : T\mathcal{D} \rightarrow \mathfrak{o}, \\ \omega &= \Lambda^{-1} \dot{\Lambda} : \mathcal{D} \rightarrow \mathfrak{o}, \\ \Gamma &= \Lambda^{-1} \mathbf{d}r : T\mathcal{D} \rightarrow E, \\ \gamma &= \Lambda^{-1} \dot{r} : \mathcal{D} \rightarrow E, \\ \rho &= \Lambda^{-1} r : \mathcal{D} \rightarrow E.\end{aligned}\tag{2.20}$$

Thus, if we interpret (Λ, r) as a curve in the group $\mathcal{F}(\mathcal{D}, \mathcal{O} \otimes E)$, the previous definition can be rewritten as

$$\begin{aligned}(\omega, \gamma) &= (\Lambda, r)^{-1}(\dot{\Lambda}, \dot{r}), \\ (\Omega, \Gamma, \rho) &= c((\Lambda, r)^{-1}),\end{aligned}$$

where c is defined by

$$c(\Lambda, r) = \left((\Lambda, r) \mathbf{d}(\Lambda, r)^{-1}, -r \right).\tag{2.21}$$

Remarkably, c is a *group one-cocycle*. Thus, the previous definition simply says that (Ω, Γ, ρ) are affine advected quantities with zero initial values. This observation strongly suggests a relation with the affine Euler–Poincaré theory developed in the context of complex fluids in [12]. In this context, the Lie group \mathcal{O} encodes the order parameter.

On the other hand, if we interpret (Λ, r) as a section of the trivial principal bundle

$$(\mathcal{D} \times \mathbb{R}) \times (\mathcal{O} \otimes E) \rightarrow \mathcal{D} \times \mathbb{R}$$

over spacetime $\mathcal{D} \times \mathbb{R}$, definition (2.20) simply says that the variables (Ω, ω, ρ) are obtained by reduction by the subgroup \mathcal{O} of the first jet extension of (Λ, r) . This, in turn, leads to a relation with the covariant Lagrange–Poincaré reduction for field theories developed in [27]. Note that by choosing the one-dimensional interval $\mathcal{D} = [0, L]$, the Lie group $\mathcal{O} = SO(3)$, and the left representation space $E = \mathbb{R}^3$, one recovers the advection of charged strands discussed earlier.

Remark 2.6. Generalizing to higher dimensions reveals certain distinct aspects of the underlying geometry of the problem that are not distinguished in considering the particular case of the charged strands. For example, in the case of *charged sheets* or *charged elastic deformed media*, \mathcal{D} is a domain in \mathbb{R}^n , with $n = 2$ or 3 , respectively, so the coordinate s has several dimensions. Then, Γ should be considered as a set of vectors $\Gamma_1, \dots, \Gamma_n$. Likewise, for the problem of flexible strands of rigid charge conformations, the distinct objects E and \mathfrak{o} both coincide with \mathbb{R}^3 . This coincidence is removed in higher dimensions and thereby clarifies the underlying geometric structure of the theory.

3. Derivation of strand dynamics in the convective representation

In this section we derive the convective equations of motion for a charged strand using a method based on the classical Euler–Poincaré (EP) approach. For expositions of the classical EP approach, see, for example, [30,19]. The EP approach is based on applying Hamilton’s variational principle to the symmetry-reduced Lagrangian and constraining the variations properly. The EP approach yields equations of motion for a nonlocally interacting strand (3.23) and (3.26) in the convective representation. Using this approach, we are also able to derive corresponding conservation laws in Section 3.2. We then proceed with the derivation of Lie–Poisson Hamiltonian formulation of strand dynamics in Section 3.3.

We explain the “bare hands” EP derivation here because we believe that doing so enhances understanding of the other sections of the paper, by providing a direct and explicit derivation of those equations of motion that invokes the Lie group action on the configuration space and thereby provides additional information.

In particular, the EP approach reveals how the Lie group action on the configuration space induces the affine structure of the EP equations (3.23) and (3.26) discussed further in Section 5. The EP approach also yields information that explains precisely how the canonical phase space (the cotangent bundle of the configuration manifold) maps to the Lie–Poisson space associated to the action, which is the dual of the Lie algebra of symmetries via the momentum map induced by the infinitesimal affine Lie algebra action. Thus, we explore the EP route in detail in this section, because it explicitly reveals the role of the Lie group action in symmetry reduction. In Section 5 it will be shown that the derivation of the EP equations and of the associated variational principle are corollaries of general theorems for systems whose configuration space is a Lie group.

A second derivation of the key formulas for strand dynamics in the convective representation is given in Appendix B. This second derivation is based on the different, but equivalent, Hamilton–Pontryagin (HP) approach in control theory (see, for example, [5]), also modified to include additional terms describing nonlocal contributions.

3.1. A modified Euler–Poincaré approach

3.1.1. Variations: Definitions Let us compute variations of ρ , ω , γ , Ω , and Γ . We proceed by first computing

$$\delta\rho = -\Lambda^{-1}\delta\Lambda\Lambda^{-1}\mathbf{r} + \Lambda^{-1}\delta\mathbf{r} = -\Sigma\rho + \Psi = -\Sigma \times \rho + \Psi = \rho \times \Sigma + \Psi, \quad (3.1)$$

where we have defined the variational quantities

$$\Sigma := \Lambda^{-1}\delta\Lambda \quad \text{and} \quad \Psi := \Lambda^{-1}\delta\mathbf{r}.$$

Next, we compute the space and time derivatives of Σ and Ψ along the curve. We have the space derivative,

$$\frac{\partial \Psi}{\partial s} = -\Lambda^{-1} \Lambda' \Lambda^{-1} \delta \mathbf{r} + \Lambda^{-1} \delta \mathbf{r}' = -\Omega \Psi + \Lambda^{-1} \delta \mathbf{r}' = -\Omega \times \Psi + \Lambda^{-1} \delta \mathbf{r}', \quad (3.2)$$

and the time derivative,

$$\frac{\partial \Psi}{\partial t} = -\Lambda^{-1} \dot{\Lambda} \Lambda^{-1} \delta \mathbf{r} + \Lambda^{-1} \delta \dot{\mathbf{r}} = -\omega \Psi + \Lambda^{-1} \delta \dot{\mathbf{r}} = -\omega \times \Psi + \Lambda^{-1} \delta \dot{\mathbf{r}}. \quad (3.3)$$

Analogously, for the space derivative of Σ ,

$$\frac{\partial \Sigma}{\partial s} = -\Lambda^{-1} \Lambda' \Lambda^{-1} \delta \Lambda + \Lambda^{-1} \delta \Lambda' = -\Omega \Sigma + \Lambda^{-1} \delta \Lambda', \quad (3.4)$$

while the time derivative of Σ is computed as follows:

$$\frac{\partial \Sigma}{\partial t} = -\Lambda^{-1} \dot{\Lambda} \Lambda^{-1} \delta \Lambda + \Lambda^{-1} \delta \dot{\Lambda} = -\omega \Sigma + \Lambda^{-1} \delta \dot{\Lambda}. \quad (3.5)$$

Now we are ready to compute the variations $\delta \gamma$, $\delta \Gamma$, $\delta \omega$, and $\delta \Omega$. The first of these is

$$\delta \gamma = -\Lambda^{-1} \delta \Lambda \Lambda^{-1} \dot{\rho} + \underbrace{\Lambda^{-1} \delta \dot{\rho}}_{\text{use (3.3)}} = -\Sigma \gamma + \omega \Psi + \frac{\partial \Psi}{\partial t},$$

so in vector form,

$$\delta \gamma = \gamma \times \Sigma + \omega \times \Psi + \frac{\partial \Psi}{\partial t}. \quad (3.6)$$

Likewise,

$$\delta \Gamma = -\Lambda^{-1} \delta \Lambda \Lambda^{-1} \rho' + \underbrace{\Lambda^{-1} \delta \rho'}_{\text{use (3.2)}} = -\Sigma \Gamma + \Omega \Psi + \frac{\partial \Psi}{\partial s},$$

which has the vector form,

$$\delta \Gamma = \Gamma \times \Sigma + \Omega \times \Psi + \frac{\partial \Psi}{\partial s}. \quad (3.7)$$

Next,

$$\delta \omega = -\Lambda^{-1} \delta \Lambda \Lambda^{-1} \dot{\Lambda} + \underbrace{\Lambda^{-1} \delta \dot{\Lambda}}_{\text{use (3.5)}} = -\Sigma \omega + \omega \Sigma + \frac{\partial \Sigma}{\partial t} = [\omega, \Sigma] + \frac{\partial \Sigma}{\partial t},$$

so expressing these formulas in terms of vectors yields

$$\delta \omega = \omega \times \Sigma + \frac{\partial \Sigma}{\partial t}. \quad (3.8)$$

Next,

$$\delta\Omega = -\Lambda^{-1}\delta\Lambda\Lambda^{-1}\Lambda' + \underbrace{\Lambda^{-1}\delta\Lambda'}_{\text{use (3.4)}} = -\Sigma\Omega + \Omega\Sigma + \frac{\partial\Sigma}{\partial s} = [\Omega, \Sigma] + \frac{\partial\Sigma}{\partial s},$$

so, again, expressing in terms of vectors leads to

$$\delta\Omega = \Omega \times \Sigma + \frac{\partial\Sigma}{\partial s}. \quad (3.9)$$

The variation of $\xi(s, s')$ is given by

$$\xi^{-1}\delta\xi(s, s') = -\text{Ad}_{\xi^{-1}(s, s')}\Sigma(s) + \Sigma(s'). \quad (3.10)$$

Finally, since $\kappa(s, s')$ is defined by

$$\kappa(s, s') = \Lambda^{-1}(s) (\mathbf{r}(s') - \mathbf{r}(s)) = \xi(s, s')\boldsymbol{\rho}(s') - \boldsymbol{\rho}(s). \quad (3.11)$$

the variation of κ has the expression

$$\begin{aligned} \delta\kappa(s, s') &= \Sigma(s)\kappa(s, s') - \Psi(s) + \xi(s, s')\Psi(s') \\ &= \Sigma(s) \times \kappa(s, s') - \Psi(s) + \xi(s, s')\Psi(s'). \end{aligned} \quad (3.12)$$

3.1.2. Derivation of the equations of motion Suppose now we want to compute variations of the reduced energy Lagrangian l which is a functional of $(\boldsymbol{\rho}, \boldsymbol{\gamma}, \boldsymbol{\Gamma}, \omega, \Omega)$. From (2.6) we see that

$$\begin{aligned} d_{k,m}(s, s') &= |\Lambda^{-1}(s)\mathbf{r}(s, t) - \Lambda^{-1}(s, t)\mathbf{r}(s', t) + \boldsymbol{\eta}_k(s) - \xi(s, s')\boldsymbol{\eta}_m(s')| \\ &= |\kappa(s, s') + \boldsymbol{\eta}_k(s) - \xi(s, s')\boldsymbol{\eta}_m(s')|. \end{aligned} \quad (3.13)$$

Let us first define the Lagrangian l as the sum of a local part l_{loc} and a nonlocal part l_{np} , according to

$$\begin{aligned} l(\omega, \boldsymbol{\gamma}, \boldsymbol{\Omega}, \boldsymbol{\Gamma}, \boldsymbol{\rho}, \xi, \kappa) &:= l_{\text{loc}} + l_{\text{np}} \\ &= l_{\text{loc}}(\omega, \boldsymbol{\gamma}, \boldsymbol{\Omega}, \boldsymbol{\Gamma}, \boldsymbol{\rho}) + \iint U(\kappa(s, s'), \xi(s, s'), \boldsymbol{\Gamma}(s), \boldsymbol{\Gamma}(s')) \, ds \, ds'. \end{aligned} \quad (3.14)$$

Note. From now on, we assume that the nonlocal part of the potential energy U is a function of the two variables $\kappa(s, s')$ and $\xi(s, s')$, as well as $\boldsymbol{\Gamma}$, since s is not necessarily the arc length. In particular, for a potential energy depending on the distance $d_{k,m}$, the variables κ and ξ enter in the linear combination defined by (3.13). In principle, the potential energy could have been chosen to be an arbitrary functional of $\Lambda^{-1}(s)\mathbf{r}(s)$, $\Lambda^{-1}(s)\mathbf{r}(s')$, and $\xi(s, s')$. Euler–Poincaré methods would be directly applicable to these functionals as well.

The equations of motion are computed from the stationary action principle $\delta S = 0$, with $S = \int l dt$ and $l = l_{\text{loc}} + l_{\text{np}}$ in equation (3.14). We have

$$\begin{aligned} \delta S = \int \left(\left\langle \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\rho}}, \delta \boldsymbol{\rho} \right\rangle + \left\langle \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\gamma}}, \delta \boldsymbol{\gamma} \right\rangle + \left\langle \frac{\delta (l_{\text{loc}} + l_{\text{np}})}{\delta \boldsymbol{\Gamma}}, \delta \boldsymbol{\Gamma} \right\rangle + \left\langle \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\omega}}, \delta \boldsymbol{\omega} \right\rangle \right. \\ \left. + \left\langle \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\Omega}}, \delta \boldsymbol{\Omega} \right\rangle + \left\langle \frac{\delta l_{\text{np}}}{\delta \boldsymbol{\kappa}}, \delta \boldsymbol{\kappa} \right\rangle + \left\langle \xi^{-1} \frac{\delta l_{\text{np}}}{\delta \xi}, \xi^{-1} \delta \xi \right\rangle \right) dt = 0, \end{aligned} \quad (3.15)$$

where $\langle \cdot, \cdot \rangle = \int (\cdot, \cdot)_{\mathbb{R}^3} ds$ represents L^2 pairing in the filament variable s . We may now substitute $\delta \boldsymbol{\rho}$ from (3.1), $\delta \boldsymbol{\gamma}$ from (3.6), and $\delta \boldsymbol{\Omega}$ from (3.9). We have

$$\left\langle \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\rho}}, \delta \boldsymbol{\rho} \right\rangle = \left\langle \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\rho}}, \boldsymbol{\rho} \times \boldsymbol{\Sigma} + \boldsymbol{\Psi} \right\rangle = \left\langle \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\rho}} \times \boldsymbol{\rho}, \boldsymbol{\Sigma} \right\rangle + \left\langle \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\rho}}, \boldsymbol{\Psi} \right\rangle. \quad (3.16)$$

For $\delta \boldsymbol{\kappa}$ we obtain

$$\begin{aligned} \left\langle \frac{\delta l_{\text{np}}}{\delta \boldsymbol{\kappa}}, \delta \boldsymbol{\kappa} \right\rangle = \int \left[\left\langle \int \frac{\partial U}{\partial \boldsymbol{\kappa}}(s, s') \times \boldsymbol{\kappa}(s, s') ds', \boldsymbol{\Sigma}(s) \right\rangle \right. \\ \left. + \left\langle \int \left(\xi(s, s') \frac{\partial U}{\partial \boldsymbol{\kappa}}(s', s) - \frac{\partial U}{\partial \boldsymbol{\kappa}}(s, s') \right) ds', \boldsymbol{\Psi}(s) \right\rangle \right] dt. \end{aligned} \quad (3.17)$$

Next,

$$\begin{aligned} & \left\langle \frac{\delta (l_{\text{loc}} + l_{\text{np}})}{\delta \boldsymbol{\Gamma}}, \delta \boldsymbol{\Gamma} \right\rangle \\ &= \left\langle \frac{\delta (l_{\text{loc}} + l_{\text{np}})}{\delta \boldsymbol{\Gamma}}, \boldsymbol{\Gamma} \times \boldsymbol{\Sigma} + \boldsymbol{\Omega} \times \boldsymbol{\Psi} + \frac{\partial \boldsymbol{\Psi}}{\partial s} \right\rangle \\ &= \left\langle \frac{\delta (l_{\text{loc}} + l_{\text{np}})}{\delta \boldsymbol{\Gamma}} \times \boldsymbol{\Gamma}, \boldsymbol{\Sigma} \right\rangle + \left\langle \frac{\delta (l_{\text{loc}} + l_{\text{np}})}{\delta \boldsymbol{\Gamma}} \times \boldsymbol{\Omega} - \frac{\partial}{\partial s} \frac{\delta (l_{\text{loc}} + l_{\text{np}})}{\delta \boldsymbol{\Gamma}}, \boldsymbol{\Psi} \right\rangle, \end{aligned} \quad (3.18)$$

and

$$\begin{aligned} \left\langle \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\gamma}}, \delta \boldsymbol{\gamma} \right\rangle &= \left\langle \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\gamma}}, \boldsymbol{\gamma} \times \boldsymbol{\Sigma} + \boldsymbol{\omega} \times \boldsymbol{\Psi} + \frac{\partial \boldsymbol{\Psi}}{\partial t} \right\rangle \\ &= \left\langle \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\gamma}} \times \boldsymbol{\gamma}, \boldsymbol{\Sigma} \right\rangle + \left\langle \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\gamma}} \times \boldsymbol{\omega} - \frac{\partial}{\partial t} \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\gamma}}, \boldsymbol{\Psi} \right\rangle. \end{aligned} \quad (3.19)$$

Variations in $\boldsymbol{\omega}$ and $\boldsymbol{\Omega}$ give, respectively, after integrating by parts,

$$\begin{aligned} \int \left\langle \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\omega}}, \delta \boldsymbol{\omega} \right\rangle dt &= \int \left\langle \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\omega}}, \boldsymbol{\omega} \times \boldsymbol{\Sigma} + \frac{\partial \boldsymbol{\Sigma}}{\partial t} \right\rangle dt \\ &= \int \left\langle \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\omega}} \times \boldsymbol{\omega} - \frac{\partial}{\partial t} \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\omega}}, \boldsymbol{\Sigma} \right\rangle dt, \end{aligned} \quad (3.20)$$

and

$$\left\langle \frac{\delta l_{\text{loc}}}{\delta \mathbf{\Omega}}, \delta \mathbf{\Omega} \right\rangle = \left\langle \frac{\delta l_{\text{loc}}}{\delta \mathbf{\Omega}}, \mathbf{\Omega} \times \mathbf{\Sigma} + \frac{\partial \mathbf{\Sigma}}{\partial s} \right\rangle = \left\langle \frac{\delta l_{\text{loc}}}{\delta \mathbf{\Omega}} \times \mathbf{\Omega} - \frac{\partial}{\partial s} \frac{\delta l_{\text{loc}}}{\delta \mathbf{\Omega}}, \mathbf{\Sigma} \right\rangle. \quad (3.21)$$

Finally, one computes the variations in ξ as follows:

$$\begin{aligned} & \int \left\langle \xi^{-1} \frac{\delta l_{\text{np}}}{\delta \xi}, \xi^{-1} \delta \xi \right\rangle ds' \\ &= \int \left\langle \xi^{-1}(s, s') \frac{\partial U}{\partial \xi}(s, s'), -\text{Ad}_{\xi^{-1}(s, s')} \Sigma(s) + \Sigma(s') \right\rangle_{\mathfrak{so}(3)} ds', \end{aligned} \quad (3.22)$$

where $\langle \cdot, \cdot \rangle_{\mathfrak{so}(3)} : \mathfrak{so}(3)^* \times \mathfrak{so}(3) \rightarrow \mathbb{R}$ is the real-valued pairing between the Lie algebra $\mathfrak{so}(3)$ and its dual $\mathfrak{so}(3)^*$.

Substitution of (3.16)–(3.22) in (3.15) gives an expression for δS that is linear in $\mathbf{\Sigma}$ and $\mathbf{\Psi}$. Collecting those terms when imposing $\delta S = 0$ implies from the term proportional to $\mathbf{\Sigma}$ that:

$$\begin{aligned} & \left(\frac{\partial}{\partial t} \frac{\delta l_{\text{loc}}}{\delta \mathbf{\omega}} + \mathbf{\omega} \times \frac{\delta l_{\text{loc}}}{\delta \mathbf{\omega}} \right) + \left(\frac{\partial}{\partial s} \frac{\delta l_{\text{loc}}}{\delta \mathbf{\Omega}} + \mathbf{\Omega} \times \frac{\delta l_{\text{loc}}}{\delta \mathbf{\Omega}} \right) \\ &= \frac{\delta l_{\text{loc}}}{\delta \mathbf{\gamma}} \times \mathbf{\gamma} + \frac{\delta (l_{\text{loc}} + l_{\text{np}})}{\delta \mathbf{\Gamma}} \times \mathbf{\Gamma} + \frac{\delta l_{\text{loc}}}{\delta \mathbf{\rho}} \times \mathbf{\rho} \\ &+ \int \left(\frac{\partial U}{\partial \mathbf{\kappa}}(s, s') \times \mathbf{\kappa}(s, s') + \mathbf{Z}(s, s') \right) ds', \end{aligned} \quad (3.23)$$

where the term $\mathbf{Z}(s, s')$ is the vector given by

$$\widehat{\mathbf{Z}}(s, s') = \xi(s, s') \left(\frac{\partial U}{\partial \xi}(s, s') \right)^T - \frac{\partial U}{\partial \xi}(s, s') \xi^T(s, s'). \quad (3.24)$$

Formula (3.24) is computed from the variation in (3.22) as follows

$$\begin{aligned} & \iint \left\langle \xi^{-1}(s, s') \frac{\partial U}{\partial \xi}(s, s'), -\text{Ad}_{\xi^{-1}(s, s')} \Sigma(s) + \Sigma(s') \right\rangle_{\mathfrak{so}(3)} ds ds' \\ &= \iint \left\langle -\text{Ad}_{\xi^{-1}(s, s')}^* \xi^T(s, s') \frac{\partial U}{\partial \xi}(s, s') + \xi^T(s', s) \frac{\partial U}{\partial \xi}(s', s), \Sigma(s) \right\rangle_{\mathfrak{so}(3)} ds ds' \\ &= \iint \left\langle -\xi(s, s') \xi^T(s, s') \frac{\partial U}{\partial \xi}(s, s') \xi^T(s, s') \right. \\ &\quad \left. + \xi(s, s') \left(\frac{\partial U}{\partial \xi}(s, s') \right)^T, \Sigma(s) \right\rangle_{\mathfrak{so}(3)} ds ds' \\ &= \iint \left\langle -\frac{\partial U}{\partial \xi}(s, s') \xi^T(s, s') + \xi(s, s') \left(\frac{\partial U}{\partial \xi}(s, s') \right)^T, \Sigma(s) \right\rangle_{\mathfrak{so}(3)} ds ds'. \end{aligned} \quad (3.25)$$

Here, we have used the fact that $\xi^T(s, s') = \xi^{-1}(s, s')$, and $\xi(s', s) = \xi^{-1}(s', s)$.

Next, we collect the terms proportional to Ψ in order to close the system. We find

$$\begin{aligned} & \left(\frac{\partial}{\partial t} \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\gamma}} + \boldsymbol{\omega} \times \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\gamma}} \right) + \left(\frac{\partial}{\partial s} \frac{\delta (l_{\text{loc}} + l_{\text{np}})}{\delta \boldsymbol{\Gamma}} + \boldsymbol{\Omega} \times \frac{\delta (l_{\text{loc}} + l_{\text{np}})}{\delta \boldsymbol{\Gamma}} \right) \\ &= \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\rho}} + \int \left(\xi(s, s') \frac{\partial U}{\partial \boldsymbol{\kappa}}(s', s) - \frac{\partial U}{\partial \boldsymbol{\kappa}}(s, s') \right) ds'. \end{aligned} \quad (3.26)$$

Summary. Equations (3.23), (3.26), (2.17), (2.18) and (2.19) represent the generalization of the Kirchhoff model that we have sought. We shall see next that under a certain transformation of variables this model reduces to a conservation law formulated in terms of the coadjoint action of $SE(3)$ on the dual of its Lie algebra $\mathfrak{se}(3)$. The same equations are derived in Appendix B using the Hamilton–Pontryagin approach, also modified to account for nonlocal dependence. This modified Hamilton–Pontryagin approach provides a direct and explicit derivation of those equations of motion that further enhances understanding of the convective representation for nonlocal dependence.

3.2. Conservation laws

In order to elucidate the physical meaning of the somewhat complex looking equations (3.23) and (3.26), we shall write them explicitly as conservation laws. For this purpose, we invoke the following identities valid for any Lie group G . Given a smooth curve $g(t) \in G$, $\eta \in \mathfrak{g}$, and $\mu \in \mathfrak{g}^*$, we have

$$\text{Ad}_{g^{-1}(t)} \frac{\partial}{\partial t} \text{Ad}_{g(t)} \eta = \text{ad}_{\sigma(t)} \eta, \quad (3.27)$$

$$\text{Ad}_{g(t)}^* \frac{\partial}{\partial t} \text{Ad}_{g^{-1}(t)}^* \mu = -\text{ad}_{\sigma(t)}^* \mu, \quad (3.28)$$

where $\sigma(t) = g^{-1} \dot{g}(t) \in \mathfrak{g}$ and Ad^* denotes the coadjoint action of G on \mathfrak{g}^* defined by $\langle \text{Ad}_g^* \mu, \eta \rangle := \langle \mu, \text{Ad}_g \eta \rangle$. Formula (3.28) generalizes to a curve $\mu(t)$ as

$$\text{Ad}_{g(t)}^* \frac{\partial}{\partial t} \text{Ad}_{g^{-1}(t)}^* \mu(t) = \dot{\mu}(t) - \text{ad}_{\sigma(t)}^* \mu(t). \quad (3.29)$$

To derive the conservation form of equations (3.23) and (3.26) we need to consider the group $G = SE(3)$ whose elements are denoted by $g = (\Lambda, \mathbf{r})$. Consider the function $(\Lambda(s, t), \mathbf{r}(s, t))$ defined on spacetime. Then we have

$$\sigma = (\Lambda, \mathbf{r})^{-1} (\dot{\Lambda}, \dot{\mathbf{r}}) = (\Lambda^{-1} \dot{\Lambda}, \Lambda^{-1} \dot{\mathbf{r}}) = (\boldsymbol{\omega}, \boldsymbol{\gamma}). \quad (3.30)$$

Recall (see, for example, [19, 30]) that the coadjoint action on $(\boldsymbol{\mu}, \boldsymbol{\beta}) \in \mathfrak{se}(3)^*$ is given by (1.14) and hence

$$\text{ad}_{(\boldsymbol{\omega}, \boldsymbol{\gamma})}^* (\boldsymbol{\mu}, \boldsymbol{\beta}) = -(\boldsymbol{\omega} \times \boldsymbol{\mu} + \boldsymbol{\gamma} \times \boldsymbol{\beta}, \boldsymbol{\omega} \times \boldsymbol{\beta}). \quad (3.31)$$

Then, using equations (3.29) and (3.31) for the temporal dual Lie algebra elements $(\boldsymbol{\mu}, \boldsymbol{\beta}) = (\delta l / \delta \boldsymbol{\omega}, \delta l / \delta \boldsymbol{\gamma})$ yields

$$\begin{aligned} & \text{Ad}_{(\Lambda, r)}^* \frac{\partial}{\partial t} \left[\text{Ad}_{(\Lambda, r)}^{*-1} \left(\frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\omega}}, \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\gamma}} \right) \right] \\ &= \frac{\partial}{\partial t} \left(\frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\omega}}, \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\gamma}} \right) + \left(\boldsymbol{\omega} \times \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\omega}} + \boldsymbol{\gamma} \times \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\gamma}}, \boldsymbol{\omega} \times \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\gamma}} \right). \end{aligned} \quad (3.32)$$

For the derivative with respect to curve parametrization s , we need to remember that the nonlocal part of the potential depends on $\boldsymbol{\Gamma}$ as well. Thus, we have

$$\begin{aligned} & \text{Ad}_{(\Lambda, r)}^* \frac{\partial}{\partial s} \left[\text{Ad}_{(\Lambda, r)}^{*-1} \left(\frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\Omega}}, \frac{\delta(l_{\text{loc}} + l_{\text{np}})}{\delta \boldsymbol{\Gamma}} \right) \right] \\ &= \frac{\partial}{\partial s} \left(\frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\Omega}}, \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\Gamma}} \right) + \left(\boldsymbol{\Omega} \times \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\Omega}} + \boldsymbol{\Gamma} \times \frac{\delta(l_{\text{loc}} + l_{\text{np}})}{\delta \boldsymbol{\Gamma}}, \boldsymbol{\Omega} \times \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\Gamma}} \right). \end{aligned} \quad (3.33)$$

Some additional identities derived below will be needed in treating the nonlocal part of the potential.

First we deal with the nonlocal term by referring to equation (3.24). This can be expressed as a formal derivative of the nonlocal part of the potential with respect to Lie algebra elements $\boldsymbol{\Omega}$ and $\boldsymbol{\Gamma}$ as follows. Note that there are only *two* free variations $\hat{\boldsymbol{\Sigma}} = \Lambda^{-1} \delta \Lambda$ and $\boldsymbol{\Psi} = \Lambda^{-1} \delta \mathbf{r}$. On the other hand, the nonlocal part of the Lagrangian depends on *three* variables $\boldsymbol{\rho}$, ξ , and $\boldsymbol{\Gamma}$. Thus, there must be a relation between the partial derivatives of the nonlocal part of the Lagrangian and the total derivatives with respect to $\boldsymbol{\Gamma}$ and $\boldsymbol{\Omega}$. This relation is computed as follows.

Upon identifying coefficients of the free variations $\boldsymbol{\Sigma} \times = \Lambda^{-1} \delta \Lambda$ and $\boldsymbol{\Psi} = \Lambda^{-1} \delta \mathbf{r}$, the following identity relates different variational derivatives of the nonlocal potential l_{np} :

$$\begin{aligned} \delta l_{\text{np}} &= \left\langle \xi^{-1} \frac{\delta l_{\text{np}}}{\delta \xi}, \xi^{-1} \delta \xi \right\rangle + \left\langle \frac{\delta l_{\text{np}}}{\delta \boldsymbol{\kappa}}, \delta \boldsymbol{\kappa} \right\rangle + \left\langle \frac{\delta l_{\text{np}}}{\delta \boldsymbol{\Gamma}}, \delta \boldsymbol{\Gamma} \right\rangle \\ &= \left\langle \frac{\delta l_{\text{np}}}{\delta \boldsymbol{\Gamma}} \Big|_{\text{Tot}}, \delta \boldsymbol{\Gamma} \right\rangle + \left\langle \frac{\delta l_{\text{np}}}{\delta \boldsymbol{\Omega}}, \delta \boldsymbol{\Omega} \right\rangle. \end{aligned} \quad (3.34)$$

We will discuss this point in detail in Section 5.3.2. Here, the subscript on $(\cdot)|_{\text{Tot}}$ denotes the *total* derivative with respect to $\boldsymbol{\Gamma}$. Using expressions (3.10) for $\xi^{-1} \delta \xi$, (3.12) for $\delta \boldsymbol{\kappa}$, (3.9) for $\delta \boldsymbol{\Omega}$, and (3.6) for $\delta \boldsymbol{\Gamma}$, then collecting terms proportional to the free variation $\boldsymbol{\Sigma}$, yields the following identity, which implicitly defines $\delta l_{\text{np}} / \delta \boldsymbol{\Omega}$ in terms of known quantities,

$$-\frac{\partial}{\partial s} \frac{\delta l_{\text{np}}}{\delta \boldsymbol{\Omega}} - \boldsymbol{\Omega} \times \frac{\delta l_{\text{np}}}{\delta \boldsymbol{\Omega}} = \int \frac{\partial U}{\partial \boldsymbol{\kappa}}(s, s') \times \boldsymbol{\kappa}(s, s') ds' + \int \mathbf{Z}(s, s') ds', \quad (3.35)$$

where we have defined $\mathbf{Z}(s, s')$ by (3.24). Likewise, identifying terms multiplying Ψ gives

$$\begin{aligned} -\frac{\partial}{\partial s} \frac{\delta l_{\text{np}}}{\delta \mathbf{\Gamma}} \Big|_{\text{Tot}} - \mathbf{\Omega} \times \frac{\delta l_{\text{np}}}{\delta \mathbf{\Gamma}} \Big|_{\text{Tot}} &= -\frac{\partial}{\partial s} \frac{\delta l_{\text{np}}}{\delta \mathbf{\Gamma}} - \mathbf{\Omega} \times \frac{\delta l_{\text{np}}}{\delta \mathbf{\Gamma}} \\ &+ \int \left(\frac{\partial U}{\partial \boldsymbol{\kappa}}(s, s') - \xi(s, s') \frac{\partial U}{\partial \boldsymbol{\kappa}}(s', s) \right) ds'. \end{aligned} \quad (3.36)$$

Therefore, we conclude that equations (3.23) and (3.26) are equivalent to the following equations expressed on $\mathfrak{se}(3)^*$ in conservative form using variations of the total Lagrangian, $l := l_{\text{loc}} + l_{\text{np}}$:

$$\begin{aligned} \frac{\partial}{\partial t} \left[\text{Ad}_{(\Lambda, r)}^* \left(\frac{\delta l}{\delta \boldsymbol{\omega}}, \frac{\delta l}{\delta \boldsymbol{\gamma}} \right) \right] + \frac{\partial}{\partial s} \left[\text{Ad}_{(\Lambda, r)}^* \left(\frac{\delta l}{\delta \mathbf{\Omega}}, \frac{\delta l}{\delta \mathbf{\Gamma}} \Big|_{\text{Tot}} \right) \right] \\ = \text{Ad}_{(\Lambda, r)}^* \left(\frac{\delta l}{\delta \boldsymbol{\rho}} \times \boldsymbol{\rho}, \frac{\delta l}{\delta \boldsymbol{\rho}} \right). \end{aligned} \quad (3.37)$$

Here, the components of

$$\text{Ad}_{(\Lambda, r)}^* \left(\frac{\delta l}{\delta \boldsymbol{\omega}}, \frac{\delta l}{\delta \boldsymbol{\gamma}} \right)$$

represent, respectively, the spatial angular momentum density and the spatial linear momentum density of the strand, whose center of mass lies on its center line. The components of

$$\text{Ad}_{(\Lambda, r)}^* \left(\frac{\delta l}{\delta \boldsymbol{\rho}} \times \boldsymbol{\rho}, \frac{\delta l}{\delta \boldsymbol{\rho}} \right) = \left(0, \Lambda \frac{\delta l}{\delta \boldsymbol{\rho}} \right)$$

are the external torques and forces. (See (1.24) for the last simplification.) As mentioned above, only external forces arising from potentials are considered in this paper. In principle, more general non-conservative forces and torques can be considered as well, but we shall leave this question for further studies.

Remark 3.1. For future reference, it is advantageous to write out the conservation law (3.37) in convective form as

$$\begin{cases} (\partial_t + \boldsymbol{\omega} \times) \frac{\delta l}{\delta \boldsymbol{\omega}} + (\partial_s + \mathbf{\Omega} \times) \frac{\delta l}{\delta \mathbf{\Omega}} + \boldsymbol{\rho} \times \frac{\delta l}{\delta \boldsymbol{\rho}} + \mathbf{\Gamma} \times \frac{\delta l}{\delta \mathbf{\Gamma}} + \boldsymbol{\gamma} \times \frac{\delta l}{\delta \boldsymbol{\gamma}} = 0, \\ (\partial_t + \boldsymbol{\omega} \times) \frac{\delta l}{\delta \boldsymbol{\gamma}} + (\partial_s + \mathbf{\Omega} \times) \frac{\delta l}{\delta \mathbf{\Gamma}} - \frac{\delta l}{\delta \boldsymbol{\rho}} = 0. \end{cases} \quad (3.38)$$

Here we have defined the total Lagrangian $l := l_{\text{loc}} + l_{\text{np}}$, and all the variational derivatives are assumed to be the *total* derivatives. Note that these equations coincide precisely with the equations for the purely elastic filaments derived in [37].

We note that the variations with respect to $\mathbf{\Omega}$ and $\mathbf{\Gamma}$ are computed implicitly in (3.35, 3.36). To actually use these equations to explicitly describe nonlocal interactions, we must expand the derivatives with respect to ξ and $\boldsymbol{\kappa}$ in (3.38). However, we emphasize again, that it is interesting that the expressions for the nonlocal interactions formally coincide with the equations for the purely elastic motion. See Section 5.3.2 for a detailed discussion of this point.

3.3. Lie–Poisson Hamiltonian structure of strand equations

It is useful to transform the Lagrangian dynamical equations into the Hamiltonian description, both to relate these equations to previous work on elastic rods and to elucidate further their mathematical structure. We start by Legendre transforming the total Lagrangian l to the Hamiltonian,

$$h(\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\Gamma}, \boldsymbol{\rho}) = \int (\boldsymbol{\pi} \cdot \boldsymbol{\omega} + \boldsymbol{\mu} \cdot \boldsymbol{\gamma}) ds - l(\boldsymbol{\omega}, \boldsymbol{\gamma}, \boldsymbol{\Omega}, \boldsymbol{\Gamma}, \boldsymbol{\rho}), \quad (3.39)$$

where $\boldsymbol{\omega}, \boldsymbol{\gamma}$ are determined from the relations $\boldsymbol{\pi} = \delta l / \delta \boldsymbol{\omega}$ and $\boldsymbol{\mu} = \delta l / \delta \boldsymbol{\gamma}$ upon assuming that l is hyperregular. Then, equations (2.17), (2.18), (2.19), and (3.38) may be expressed in Lie–Poisson form with three cocycles as

$$\frac{\partial}{\partial t} \begin{bmatrix} \boldsymbol{\pi} \\ \boldsymbol{\mu} \\ \boldsymbol{\Omega} \\ \boldsymbol{\Gamma} \\ \boldsymbol{\rho} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\pi} \times & \boldsymbol{\mu} \times & (\partial_s + \boldsymbol{\Omega} \times) & \boldsymbol{\Gamma} \times & \boldsymbol{\rho} \times \\ \boldsymbol{\mu} \times & 0 & 0 & (\partial_s + \boldsymbol{\Omega} \times) & -\text{Id} \\ (\partial_s + \boldsymbol{\Omega} \times) & 0 & 0 & 0 & 0 \\ \boldsymbol{\Gamma} \times & (\partial_s + \boldsymbol{\Omega} \times) & 0 & 0 & 0 \\ \boldsymbol{\rho} \times & \text{Id} & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta h / \delta \boldsymbol{\pi} \\ \delta h / \delta \boldsymbol{\mu} \\ \delta h / \delta \boldsymbol{\Omega} \\ \delta h / \delta \boldsymbol{\Gamma} \\ \delta h / \delta \boldsymbol{\rho} \end{bmatrix}. \quad (3.40)$$

Note that $\boldsymbol{\omega} = \delta h / \delta \boldsymbol{\pi}$ and $\boldsymbol{\gamma} = \delta h / \delta \boldsymbol{\mu}$. The affine terms ∂_s and Id arise from the cocycle appearing in the definition of the variables $\boldsymbol{\Omega}, \boldsymbol{\Gamma}, \boldsymbol{\rho}$ in (1.1); see also (2.20). These equations produce the affine terms located in the matrix elements $\{\boldsymbol{\pi}, \boldsymbol{\Omega}\}$, $\{\boldsymbol{\mu}, \boldsymbol{\Gamma}\}$, and $\{\boldsymbol{\mu}, \boldsymbol{\rho}\}$.

This Hamiltonian matrix defines an *affine Lie–Poisson bracket* on the dual of the semidirect product Lie algebra

$$\mathcal{F}(I, \mathfrak{se}(3)) \mathbin{\mathbb{S}} \mathcal{F}(I, \mathfrak{se}(3) \times \mathbb{R}^3),$$

where $\mathfrak{se}(3) = \mathfrak{so}(3) \mathbin{\mathbb{S}} \mathbb{R}^3$, $I = [0, L]$, and

$$(\boldsymbol{\pi}, \boldsymbol{\mu}) \in \mathcal{F}(I, \mathfrak{se}(3))^* \quad \text{and} \quad (\boldsymbol{\Omega}, \boldsymbol{\Gamma}, \boldsymbol{\rho}) \in \mathcal{F}(I, \mathfrak{se}(3) \times \mathbb{R}^3)^*.$$

The associated affine Lie–Poisson bracket reads

$$\begin{aligned} \{f, g\}(\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\Gamma}, \boldsymbol{\rho}) = & - \int \boldsymbol{\pi} \cdot \left(\frac{\delta f}{\delta \boldsymbol{\pi}} \times \frac{\delta g}{\delta \boldsymbol{\pi}} \right) \\ & - \int \boldsymbol{\mu} \cdot \left(\frac{\delta f}{\delta \boldsymbol{\mu}} \times \frac{\delta g}{\delta \boldsymbol{\pi}} - \frac{\delta g}{\delta \boldsymbol{\mu}} \times \frac{\delta f}{\delta \boldsymbol{\pi}} \right) \\ & - \int \boldsymbol{\Omega} \cdot \left(\frac{\delta f}{\delta \boldsymbol{\Omega}} \times \frac{\delta g}{\delta \boldsymbol{\pi}} - \frac{\delta g}{\delta \boldsymbol{\Omega}} \times \frac{\delta f}{\delta \boldsymbol{\pi}} \right) \\ & - \int \boldsymbol{\Omega} \cdot \left(\frac{\delta f}{\delta \boldsymbol{\Gamma}} \times \frac{\delta g}{\delta \boldsymbol{\mu}} - \frac{\delta g}{\delta \boldsymbol{\Gamma}} \times \frac{\delta f}{\delta \boldsymbol{\mu}} \right) \\ & - \int \boldsymbol{\Gamma} \cdot \left(\frac{\delta f}{\delta \boldsymbol{\Gamma}} \times \frac{\delta g}{\delta \boldsymbol{\pi}} - \frac{\delta g}{\delta \boldsymbol{\Gamma}} \times \frac{\delta f}{\delta \boldsymbol{\pi}} \right) \end{aligned}$$

$$\begin{aligned}
& - \int \boldsymbol{\rho} \cdot \left(\frac{\delta f}{\delta \boldsymbol{\rho}} \times \frac{\delta g}{\delta \boldsymbol{\pi}} - \frac{\delta g}{\delta \boldsymbol{\rho}} \times \frac{\delta f}{\delta \boldsymbol{\pi}} \right) \\
& + \int \frac{\delta f}{\delta \boldsymbol{\Omega}} \cdot \partial_s \frac{\delta g}{\delta \boldsymbol{\pi}} + \frac{\delta f}{\delta \boldsymbol{\Gamma}} \cdot \partial_s \frac{\delta g}{\delta \boldsymbol{\mu}} + \frac{\delta f}{\delta \boldsymbol{\rho}} \cdot \frac{\delta g}{\delta \boldsymbol{\mu}} \\
& - \int \frac{\delta g}{\delta \boldsymbol{\Omega}} \cdot \partial_s \frac{\delta f}{\delta \boldsymbol{\pi}} + \frac{\delta g}{\delta \boldsymbol{\Gamma}} \cdot \partial_s \frac{\delta f}{\delta \boldsymbol{\mu}} + \frac{\delta g}{\delta \boldsymbol{\rho}} \cdot \frac{\delta f}{\delta \boldsymbol{\mu}}. \quad (3.41)
\end{aligned}$$

The first line represents the Lie–Poisson bracket on the dual of the Lie algebra $\mathcal{F}(I, \mathfrak{se}(3))$. The first five lines represent the Lie–Poisson bracket on the dual of the semidirect product Lie algebra

$$\mathcal{F}(I, \mathfrak{se}(3)) \oplus \mathcal{F}(I, \mathfrak{se}(3) \times \mathbb{R}^3).$$

The last two lines represent the affine terms due to the presence of a cocycle, as well as the canonical Poisson bracket in $(\boldsymbol{\rho}, \boldsymbol{\mu})$. The Poisson bracket (3.41) is an extension that includes $\boldsymbol{\rho}$ in the Poisson bracket for the exact geometric rod theory of [37] in the convective representation. Remarkably, from a geometric point of view, this Hamiltonian structure is *identical* to that of complex fluids [12, 18]. The reason for this will be explained in detail in Section 5.

4. Introduction to affine Lie group actions

Classical Lagrangian approach to the molecular strand. The Euler–Poincaré method used in Section 3 and the equivalent alternative Hamilton–Pontryagin method used in Appendix B to derive the strand equations of motion (3.23) and (3.26) strongly suggest that the dynamics of the molecular strand with nonlocal interactions can be obtained by a *classical* Lagrangian reduction. In other words, the evolution for the Lagrangian variables Λ and \mathbf{r} should be given by the *standard Euler–Lagrange equations*

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\Lambda}} - \frac{\partial L}{\partial \Lambda} = 0, \quad \frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{r}}} - \frac{\partial L}{\partial \mathbf{r}} = 0$$

associated with a Lagrangian L defined on the tangent bundle TQ of the configuration space Q and invariant under the action of the symmetry group of the theory. The evolution of the reduced quantities $\boldsymbol{\omega}, \boldsymbol{\gamma}, \boldsymbol{\Omega}, \boldsymbol{\Gamma}, \boldsymbol{\rho}$ can then be obtained by applying the general tools of Lagrangian reduction. Such an approach, together with its Hamiltonian counterpart, has been successfully applied to a wide range of mechanical systems with symmetry, from fluid dynamics and imaging to rigid bodies and particles with broken symmetries. It is therefore of great interest to obtain such a description for the molecular strand. This is the main goal of Section 5.

Perhaps not surprisingly, the description of such a complex system as the molecular strand with nonlocal interactions needs a somewhat sophisticated version of the classical Lagrangian reduction. This is explained in the present section by building up to the geometric setting via simpler examples. Besides the fact that the configuration space

$$(\Lambda, \mathbf{r}) \in \mathcal{F}([0, L], SE(3))$$

is infinite dimensional, there are two major difficulties to overcome. The first is related to the observation that the advected variables (Ω, Γ, ρ) are not acted on linearly by the Lagrangian variables (Λ, \mathbf{r}) ; see (1.1). In order to understand their evolution, an affine action (that is, a cocycle) needs to be introduced. The second difficulty is related to the nonlocal dependence of the reduced Lagrangian on the unreduced variables Λ, \mathbf{r} . Remarkably, the geometrization of this nonlocal dependence is also solved by the presence of the affine term (the cocycle) in the action, which allows the choice of the zero value for the reference condition $(\Omega_{\text{ref}}, \Gamma_{\text{ref}}, \rho_{\text{ref}})$ without leading to trivial dynamics.

Main goals of the Lagrangian approach. Besides these difficulties, such a *classical* geometric description of the strand has many advantages. First, at a pure mathematical level, it justifies rigorously the two “bare hand” derivations of the equations described in Section 3 and Appendix B from the Euler–Poincaré and Hamilton–Pontryagin approaches, respectively. The geometric description also explains the somewhat mysterious vanishing of the explicit dependence on the nonlocal terms in the final equation of motion; see (3.38). On the Hamiltonian side, the framework justifies the presence of the cocycles in the Poisson bracket (3.41) (see also (3.40)) and produces the Hamiltonian structure in the convective representation by reduction of the *canonical* Hamilton equations on the phase space $T^*\mathcal{F}([0, L], SE(3))$ of the system.

At a more applied level, this symmetry-reduced Lagrangian approach provides a guide towards a generalization to higher dimensional versions or to other matrix Lie groups describing the charge conformation. Moreover, this classical approach is also suitable if one wishes to couple the molecular strand with fluid dynamics, since the two systems are now described by the same simple geometric framework: *canonical Lagrangian (respectively Hamiltonian) description on the tangent (respectively cotangent) bundle of the configuration space.*

Pedagogical examples. We begin with three preparatory examples that illustrate the main idea behind the classical Lagrangian approach in a simpler setting than that required for the molecular strand.

Example 1. The simplest setting of Lagrangian reduction is that of a G -invariant Lagrangian $L : TG \rightarrow \mathbb{R}$, defined on the tangent bundle of its symmetry Lie group G . In this case, the Euler–Lagrange equations TG can be reduced to provide equations for an unknown in the Lie algebra \mathfrak{g} , known as the *Euler–Poincaré equations*,

$$\frac{d}{dt} \frac{\delta l}{\delta \xi} = \text{ad}_\xi^* \frac{\delta l}{\delta \xi}, \quad \xi = g^{-1} \dot{g} \in \mathfrak{g}, \quad (4.1)$$

where $l(g^{-1} \dot{g}) = L(g, \dot{g})$ is the reduced Lagrangian ; see, for example, [30]. Even such a simple setting has many applications. For example, Euler–Poincaré (EP) dynamics on the orthogonal group $G = SO(3)$ produces Euler’s rigid body equations, while EP dynamics on the group of volume preserving diffeomorphisms yields the Euler equations for ideal fluid flows.

Example 2. When passing from the free rigid body to the heavy top, the direction of gravity breaks the $SO(3)$ -invariance of the Lagrangian. In the framework of Euler–Poincaré theory, this is understood as follows. One starts with a G -invariant Lagrangian $L = L(g, \dot{g}, a) : TG \times V^* \rightarrow \mathbb{R}$, where the (dual) vector space V^* contains the advected quantity a . By fixing a particular reference value $a_{\text{ref}} \in V^*$, one breaks the symmetry and produces a physical Lagrangian $L_{a_{\text{ref}}} : TG \rightarrow \mathbb{R}$ that is only $G_{a_{\text{ref}}}$ -invariant. Here $G_{a_{\text{ref}}}$ denotes the isotropy group of the parameter a_{ref} .

For the heavy top, the parameter corresponds to the choice of a fixed direction for gravity, whereas for compressible hydrodynamics, this choice corresponds to fixing the mass density ρ_{ref} of the fluid in the reference configuration for the Lagrangian representation. In the convective picture (for the heavy top) and Eulerian picture (for fluids), the parameters are *linearly* advected by the flow of the Euler–Lagrange equation. For example, in the case of compressible fluids, the advection relation is the continuity equation

$$\dot{\rho} + \text{div}(\rho u) = 0,$$

for the mass density. As has been seen, such a linear evolution does not appear in the molecular strand and affine advection needs to be considered.

Returning to the abstract formulation, the G -invariant function $L = L(g, \dot{g}, a_{\text{ref}})$ determines the reduced Lagrangian $l(\xi, a) = L(g^{-1}\dot{g}, g^{-1}a_{\text{ref}})$ on the space $\mathfrak{g} \times V^*$, and the presence of the new variable a acted on by G modifies the right-hand side of the Euler–Poincaré equation (4.1). We refer to [21] for the theory of Euler–Poincaré reduction with advected variables.

Example 3. In order to extend classical Lagrangian and Hamiltonian reductions to the case of fluids with internal structure, such as superfluids or spin glasses, one needs to consider quantities that are *affinely advected* by the Lagrangian flow, as opposed to the *linearly* advected quantities of Example 2. Such an observation is made in [12] in order to explain, from a canonical point of view (that is, by reduction of the canonical Hamilton equations on phase space), the presence of cocycles in the Hamiltonian structure of these equations (see also [20]).

Roughly speaking, the main idea is to replace the linear action in Example 2 by an affine action. Such an action is of the form $a \mapsto ga + c(g)$, where $c(g)$ is a group one-cocycle. As before, we start with a G -invariant function $L = L(g, \dot{g}, a_{\text{ref}}) : TG \times V^* \rightarrow \mathbb{R}$ where V^* is the space of affine advected quantities. Fixing a_{ref} produces a Lagrangian $L_{a_{\text{ref}}}$ that is only invariant under the isotropy group $G_{a_{\text{ref}}}$ of the affine action at a_{ref} . As we shall see, in contrast with Example 2 above, it makes sense (and is useful) to chose the trivial initial condition $a_{\text{ref}} = 0$.

The strand with nonlocal interactions. In the next section, it is shown that in order to obtain the classical Lagrangian formulation for the molecular strand with nonlocal interactions, a slight generalization of Example 3 is needed.

The main difference is that now the $G_{a_{\text{ref}}}$ -invariant Lagrangian $L_{a_{\text{ref}}} : TG \rightarrow \mathbb{R}$ of the system is only given for the particular value $a_{\text{ref}} = 0$ of the parameter. Therefore, the reduced equations of motion take place on the space $\mathfrak{g} \times \mathcal{O}$, where $\mathcal{O} \subset V^*$ is the orbit of $a_{\text{ref}} = 0$ under the affine action of G . The somewhat mysterious link

between the nonlocal character of the Lagrangian and the fact that it is only expressible at the particular value $a_{\text{ref}} = 0$, is explained in Remark 5.7 with the help of Remark 5.2.

The associated Lagrangian reduction is called below “affine reduction at a fixed parameter” and is studied in detail in Section 5.3. For application to the strand equations, the configuration Lie group is evidently given by

$$G = \mathcal{F}([0, L], SE(3)) \ni g = (\Lambda, \mathbf{r}),$$

whereas the affine advected quantities are represented by the variables (Ω, Γ, ρ) .

When the nonlocal interactions are neglected, then the dynamics of the Kirchhoff rod in the convective representation are recovered and the situation of Example 3 re-emerges. This simpler case and its link to Lagrange–Poincaré reduction and Clebsch-constrained variational principles are explored in [11].

5. The affine Euler–Poincaré and Lie–Poisson approaches

This section explains how the equations of the charged strand may be obtained by classical Lagrangian or Hamiltonian reduction by the symmetry group. In particular, in the material representation, the dynamics is governed by the standard Euler–Lagrange or canonical Hamilton equations on the (co)tangent bundle of the configuration group $\mathcal{F}(I, SE(3))$. To see this, we shall use the process of affine Euler–Poincaré and affine Lie–Poisson reduction. This proves that the charged strand admits the same geometrical description as for complex fluids and spin systems.

We begin by recalling from [12] the theory of *affine Euler–Poincaré* and *Lie–Poisson reduction*. In contrast to [12], however, we consider here Lagrangians and Hamiltonians that are *left*-invariant, rather than *right*-invariant.

5.1. Notations for semidirect products

Let V be a vector space and assume that the Lie group G acts on the *left* by linear maps (and hence G also acts linearly on the left on the dual space V^*). As a set, the semidirect product $S = G \ltimes V$ is the Cartesian product $S = G \times V$ whose group multiplication is given by

$$(g_1, v_1)(g_2, v_2) := (g_1 g_2, v_1 + g_1 v_2),$$

where the action of $g \in G$ on $v \in V$ is denoted simply as gv . The identity element of $G \ltimes V$ is $(e, 0)$, where e is the identity element of G , and $(g, v)^{-1} = (g^{-1}, -g^{-1}v)$ for all $g \in G, v \in V$. The Lie algebra of S is the semidirect product Lie algebra, $\mathfrak{s} = \mathfrak{g} \ltimes V$, whose Lie bracket has the expression

$$\text{ad}_{(\xi_1, v_1)}(\xi_2, v_2) := [(\xi_1, v_1), (\xi_2, v_2)] = ([\xi_1, \xi_2], \xi_1 v_2 - \xi_2 v_1),$$

where ξv denotes the induced linear \mathfrak{g} -action on V , that is,

$$\xi v := \left. \frac{d}{dt} \right|_{t=0} \exp(t\xi)v \in V.$$

From the expression for the Lie bracket, it follows that for $(\xi, v) \in \mathfrak{s}$ and $(\mu, a) \in \mathfrak{s}^*$, the dual of the Lie algebra \mathfrak{s} , we have

$$\mathrm{ad}_{(\xi, v)}^*(\mu, a) = (\mathrm{ad}_\xi^* \mu - v \diamond a, -\xi a),$$

where $\xi a \in V^*$ and $v \diamond a \in \mathfrak{g}^*$ are given by

$$\xi a := \left. \frac{d}{dt} \right|_{t=0} \exp(t\xi)a \quad \text{and} \quad \langle v \diamond a, \xi \rangle_{\mathfrak{g}} := -\langle \xi a, v \rangle_V,$$

and where $\langle \cdot, \cdot \rangle_{\mathfrak{g}} : \mathfrak{g}^* \times \mathfrak{g} \rightarrow \mathbb{R}$ and $\langle \cdot, \cdot \rangle_V : V^* \times V \rightarrow \mathbb{R}$ are the duality pairings. The coadjoint action of S on \mathfrak{s}^* has the expression

$$\mathrm{Ad}_{(g, v)^{-1}}^*(\mu, a) = \left(\mathrm{Ad}_{g^{-1}}^* \mu + v \diamond ga, ga \right). \quad (5.1)$$

Suppose we are given a *left* representation of G on the vector space V^* . We can form an *affine left* representation $\theta_g(a) := ga + c(g)$, where $c \in \mathcal{F}(G, V^*)$ is a *left group one-cocycle*, that is, it verifies the property

$$c(gh) = c(g) + gc(h), \quad (5.2)$$

for all $g, h \in G$. Note that

$$\left. \frac{d}{dt} \right|_{t=0} \theta_{\exp(t\xi)}(a) = \xi a + \mathbf{d}c(\xi)$$

and

$$\langle \xi a + \mathbf{d}c(\xi), v \rangle_V = \langle \mathbf{d}c^T(v) - v \diamond a, \xi \rangle_{\mathfrak{g}},$$

where $\mathbf{d}c : \mathfrak{g} \rightarrow V^*$ is defined by $\mathbf{d}c(\xi) := T_e c(\xi)$, and $\mathbf{d}c^T : V \rightarrow \mathfrak{g}^*$ is defined by

$$\langle \mathbf{d}c^T(v), \xi \rangle_{\mathfrak{g}} := \langle \mathbf{d}c(\xi), v \rangle_V.$$

5.2. Affine Lagrangian and Hamiltonian semidirect product theory

Concerning the Lagrangian side, the general setup is the following.

- Assume that we have a function $L : TG \times V^* \rightarrow \mathbb{R}$ which is *left* G -invariant under the affine action $(v_h, a) \mapsto (gv_h, \theta_g(a)) = (gv_h, ga + c(g))$, where $g, h \in G, v_h \in TG, a \in V^*$.
- In particular, if $a_{\mathrm{ref}} \in V^*$, define the Lagrangian $L_{a_{\mathrm{ref}}} : TG \rightarrow \mathbb{R}$ by $L_{a_{\mathrm{ref}}}(v_g) := L(v_g, a_{\mathrm{ref}})$. Then $L_{a_{\mathrm{ref}}}$ is left invariant under the lift to TG of the left action of $G_{a_{\mathrm{ref}}}^c$ on G , where $G_{a_{\mathrm{ref}}}^c := \{g \in G \mid \theta_g(a_{\mathrm{ref}}) = ga_{\mathrm{ref}} + c(g) = a_{\mathrm{ref}}\}$ is the isotropy group of a_{ref} with respect to the affine action θ .
- Define $l : \mathfrak{g} \times V^* \rightarrow \mathbb{R}$ by $l := L|_{\mathfrak{g} \times V^*}$. Left G -invariance of L yields

$$l(g^{-1}v_g, \theta_{g^{-1}}(a)) = L(v_g, a)$$

for all $g \in G, v_g \in T_g G, a \in V^*$.

- For a curve $g(t) \in G$, let $\xi(t) := g(t)^{-1}\dot{g}(t) \in \mathfrak{g}$ and define the curve $a(t) \in V^*$ as the unique solution of the following affine differential equation with time dependent coefficients

$$\dot{a} = -\xi a - \mathbf{d}c(\xi), \quad (5.3)$$

with initial condition

$$a(0) = g(0)^{-1}a_{\text{ref}} + c(g(0)^{-1}) \quad \text{for } g(0) \in G. \quad (5.4)$$

The solution of (5.3) can then be written as

$$a(t) = \theta_{g(t)^{-1}}(a_{\text{ref}}) = g(t)^{-1}a_{\text{ref}} + c(g(t)^{-1}). \quad (5.5)$$

For example, the choice $a_{\text{ref}} = 0$ in one dimension means that the filament coordinates have no singularities. In an application to a two-dimensional sheet in \mathbb{R}^3 the choice $a_{\text{ref}} = 0$ would mean that no defects (disclinations) appear in the reference configuration. On such a sheet, the curvature for a choice $a_{\text{ref}} \neq 0$ would represent the areal density of the defects in the order parameter of the reference configuration. This matter has been discussed earlier for complex fluids in [12]. The present paper does not delve into defects and other higher dimensional issues.

Theorem 5.1. *In the preceding notation, the following are equivalent:*

- i With $a_{\text{ref}} \in V^*$ held fixed, Hamilton's variational principle

$$\delta \int_{t_0}^{t_1} L_{a_{\text{ref}}}(g, \dot{g}) dt = 0, \quad (5.6)$$

holds, for variations $\delta g(t)$ of $g(t)$ vanishing at the endpoints.

- ii $g(t)$ satisfies the Euler–Lagrange equations for $L_{a_{\text{ref}}}$ on G .
- iii The constrained variational principle

$$\delta \int_{t_0}^{t_1} l(\xi, a) dt = 0, \quad (5.7)$$

holds on $\mathfrak{g} \times V^$, upon using variations of the form*

$$\delta \xi = \frac{\partial \eta}{\partial t} + [\xi, \eta], \quad \delta a = -\eta a - \mathbf{d}c(\eta),$$

where $\eta(t) \in \mathfrak{g}$ vanishes at the endpoints.

- iv The affine Euler–Poincaré equations hold on $\mathfrak{g} \times V^*$:

$$\frac{\partial}{\partial t} \frac{\delta l}{\delta \xi} = \text{ad}_\xi^* \frac{\delta l}{\delta \xi} + \frac{\delta l}{\delta a} \diamond a - \mathbf{d}c^T \left(\frac{\delta l}{\delta a} \right). \quad (5.8)$$

See [12] for the proof and applications to *spin systems* and *complex fluids*. Concerning the Hamiltonian side, the setup is the following.

- Assume that we have a function $H : T^*G \times V^* \rightarrow \mathbb{R}$ which is *left* invariant under the affine action $(\alpha_h, a) \mapsto (g\alpha_h, \theta_g(a))$, for all $g, h \in G, \alpha_h \in T^*G, a \in V^*$.

– In particular, if $a_{\text{ref}} \in V^*$, define the Hamiltonian $H_{a_{\text{ref}}} : T^*G \rightarrow \mathbb{R}$ by

$$H_{a_{\text{ref}}}(\alpha_g) := H(\alpha_g, a_{\text{ref}}).$$

Then $H_{a_{\text{ref}}}$ is left invariant under the lift to T^*G of the left action of $G_{a_{\text{ref}}}^c$ on G .
 – Define $h : \mathfrak{g}^* \times V^* \rightarrow \mathbb{R}$ by $h := H|_{\mathfrak{g}^* \times V^*}$. Left G -invariance of H yields

$$h(g^{-1}\alpha_g, \theta_{g^{-1}}(a)) = H(\alpha_g, a).$$

for all $g \in G, \alpha_g \in T_g^*G, a \in V^*$.

Note that the G -action on $T^*G \times V^*$ is induced by the S -action on $T^*S = T^*G \times (V \times V^*)$ given by

$$\Psi_{(g,v)}(\alpha_h, (u, a)) := (g\alpha_h, v + gu, ga + c(g)). \quad (5.9)$$

The affine action Ψ appears as a modification of the cotangent lift of left translation on S by an affine term. Thus, we can think of the Hamiltonian $H : T^*G \times V^* \rightarrow \mathbb{R}$ as being the Poisson reduction of a S -invariant Hamiltonian $\bar{H} : T^*S \rightarrow \mathbb{R}$ by the normal subgroup $\{e\} \times V$ since $(T^*S)/(\{e\} \times V) \cong T^*G \times V^*$. Note also that every Hamiltonian $\bar{H} = \bar{H}(\alpha_h, (u, a))$, defined on T^*S and left invariant under the affine action Ψ , does not depend on the variable $u \in V$.

Theorem 5.2. *Let $\alpha(t) \in T_{g(t)}^*G$ be a solution of Hamilton's equations associated to $H_{a_{\text{ref}}}$ with initial condition $\mu_0 \in T_e^*G$. Then $(\mu(t), a(t)) := (g(t)^{-1}\alpha(t), \theta_{g(t)^{-1}}(a_{\text{ref}})) \in \mathfrak{g}^* \times V^*$ is a solution of the affine Lie–Poisson equations on \mathfrak{s}^**

$$\frac{\partial}{\partial t}(\mu, a) = \left(\text{ad}_{\frac{\delta h}{\delta \mu}}^* \mu - \frac{\delta h}{\delta a} \diamond a + \mathbf{d}c^T \left(\frac{\delta h}{\delta a} \right), -\frac{\delta h}{\delta \mu} a - \mathbf{d}c \left(\frac{\delta h}{\delta \mu} \right) \right)$$

with initial conditions $(\mu(0), a(0)) = (\mu_0, a_{\text{ref}})$. The associated Poisson bracket is the affine Lie–Poisson bracket on the dual \mathfrak{s}^*

$$\begin{aligned} \{f, g\}(\mu, a) = & - \left\langle \mu, \left[\frac{\delta f}{\delta \mu}, \frac{\delta g}{\delta \mu} \right] \right\rangle - \left\langle a, \frac{\delta f}{\delta \mu} \frac{\delta g}{\delta a} - \frac{\delta g}{\delta \mu} \frac{\delta f}{\delta a} \right\rangle \\ & + \left\langle \mathbf{d}c \left(\frac{\delta f}{\delta \mu} \right), \frac{\delta g}{\delta a} \right\rangle - \left\langle \mathbf{d}c \left(\frac{\delta g}{\delta \mu} \right), \frac{\delta f}{\delta a} \right\rangle. \end{aligned} \quad (5.10)$$

Conversely, given $\mu_0 \in T_e^*G$, the solution $\alpha(t)$ of the Hamiltonian system associated to $H_{a_{\text{ref}}}$ is reconstructed from the solution $(\mu(t), a(t))$ of the affine Lie–Poisson equations with initial conditions $(\mu(0), a(0)) = (\mu_0, a_{\text{ref}})$ by setting $\alpha(t) = g(t)\mu(t)$, where $g(t)$ is the unique solution of the differential equation $\dot{g}(t) = g(t)\frac{\delta h}{\delta \mu(t)}$ with initial condition $g(0) = e$.

See [12] for the proof of this theorem and some applications.

Remark 5.1. (Momentum maps) We now comment on the momentum maps at each stage of the reduction process. In [12] it is shown that the momentum map associated to the affine action (5.9) is given by

$$\mathbf{J} : T^*S \rightarrow \mathfrak{s}^*, \quad \mathbf{J}(\alpha_g, (u, b)) = (\alpha_g g^{-1} + u \diamond b - \mathbf{d}c^T(u), b), \quad (5.11)$$

where $g \in G, \alpha_g \in T^*G, u \in V$, and $b \in V^*$. The proof of this formula uses the general expression for the momentum map on a magnetic cotangent bundle with respect to the cotangent-lifted action. Conservation of \mathbf{J} then implies that the motion takes place on affine coadjoint orbits.

5.3. Affine reduction at fixed parameter

As we shall see, the affine reduction theorems recalled above do not apply directly to the molecular strand. This is because the Lagrangian of the molecular strand is only given for the particular value $a_{\text{ref}} = 0$ of the parameter and we do not have a concrete expression for $L_{a_{\text{ref}}}$ when $a_{\text{ref}} \neq 0$ is an arbitrary element of V^* . Extending L_0 by G -invariance only yields a Lagrangian on $TG \times \mathcal{O}_0^c$, where $\mathcal{O}_0^c \subset V^*$ is the orbit of the affine G -action on V^* . Fortunately, the Lagrangian L_0 for the molecular strand is invariant under the isotropy group $G_0^c = \{g \in G \mid c(g) = 0\}$ and this turns out to be enough for the extension of the affine semidirect product reduction theorem.

5.3.1. Lagrangian approach We consider here the case of a $G_{a_{\text{ref}}}^c$ -invariant Lagrangian $L_{a_{\text{ref}}} : TG \rightarrow \mathbb{R}$ for a fixed $a_{\text{ref}} \in V^*$, but we do not suppose that this Lagrangian comes from a G -invariant function $L : TG \times V^* \rightarrow \mathbb{R}$. In particular, we do not know the expression of L_a when $a \neq a_{\text{ref}}$ is an arbitrary element of V^* . To $L_{a_{\text{ref}}}$ we associate the reduced Lagrangian l defined on the submanifold

$$\mathfrak{g} \times \mathcal{O}_{a_{\text{ref}}}^c \subset \mathfrak{g} \times V^*, \quad \mathcal{O}_{a_{\text{ref}}}^c := \{\theta_g(a_{\text{ref}}) \mid g \in G\}$$

given by $l(\xi, \theta_g(a_{\text{ref}})) = L_{a_{\text{ref}}}(g^{-1}\xi)$. The tangent space at a to $\mathcal{O}_{a_{\text{ref}}}^c$ is given by

$$T_a \mathcal{O}_{a_{\text{ref}}}^c = \{\mathbf{d}c(\eta) + \eta a \mid \eta \in \mathfrak{g}\}. \quad (5.12)$$

The analogue of Theorem 5.1 in this case is given below.

Theorem 5.3. *Let a_{ref} be a fixed element in V^* and $g(t)$ be a curve in G with $g(0) = e$. Define the curves $\xi(t) = g(t)^{-1}\dot{g}(t) \in \mathfrak{g}$ and $a(t) := \theta_{g(t)^{-1}}a_{\text{ref}} \in V^*$. Then the following are equivalent.*

i *With a_{ref} held fixed, Hamilton's variational principle*

$$\delta \int_{t_0}^{t_1} L_{a_{\text{ref}}}(g, \dot{g}) dt = 0 \quad (5.13)$$

holds for variations $\delta g(t)$ of $g(t)$ vanishing at the endpoints.

ii *$g(t)$ satisfies the Euler–Lagrange equations for $L_{a_{\text{ref}}}$ on G .*

iii *The constrained variational principle*

$$\delta \int_{t_0}^{t_1} l(\xi, a) dt = 0 \quad (5.14)$$

holds on $\mathfrak{g} \times \mathcal{O}_{a_{\text{ref}}}^c \subset \mathfrak{g} \times V^*$, upon using variations of the form

$$\delta \xi = \frac{\partial \eta}{\partial t} + [\xi, \eta], \quad \delta a = -\eta a - \mathbf{d}c(\eta),$$

where $\eta(t) \in \mathfrak{g}$ vanishes at the endpoints.

iv Extending l arbitrarily to $\mathfrak{g} \times V^*$, the affine Euler–Poincaré equations hold on the submanifold $\mathfrak{g} \times \mathcal{O}_{a_{\text{ref}}}^c \subset \mathfrak{g} \times V^*$:

$$\frac{\partial}{\partial t} \frac{\delta l}{\delta \xi} = \text{ad}_\xi^* \frac{\delta l}{\delta \xi} + \frac{\delta l}{\delta a} \diamond a - \mathbf{d}c^T \left(\frac{\delta l}{\delta a} \right). \quad (5.15)$$

Proof. The equivalence of **i** and **ii** is true in general. The equivalence of **i** and **iii** and the equivalence of **iii** and **iv** can be shown exactly as in the standard case, that is, the case when l is defined on the whole space $\mathfrak{g} \times V^*$. The only minor difference occurs when l is differentiated with respect to the second variable. In this case the functional derivative $\delta l / \delta a \in V$ is replaced by the tangent map $\mathbf{d}_2 l(\xi, a) \in T_a^* \mathcal{O}_{a_{\text{ref}}}^c$ and one observes that

$$\mathbf{d}_2 l(\xi, a) \cdot \delta a = \left\langle \frac{\delta \tilde{l}}{\delta a}, \delta a \right\rangle, \quad \text{for all } \delta a \in T_a \mathcal{O}_{a_{\text{ref}}}^c$$

for any extension \tilde{l} of l to $\mathfrak{g} \times V^*$. Note that $\delta a = -\eta a - \mathbf{d}c(\eta) \in T_a \mathcal{O}_{a_{\text{ref}}}^c$ for $\eta \in \eta$ and that any vector in $T_a \mathcal{O}_{a_{\text{ref}}}^c$ is of this form. From now on we denote also by l , instead of \tilde{l} , an arbitrary extension of l . \square

Remark 5.2. (The case $a_{\text{ref}} = 0$ and the charged strand) For the charged molecular strand we shall need to choose $a_{\text{ref}} = 0$. In this case the isotropy group is $G_0^c = \{g \in G \mid c(g) = 0\}$. Given a G_0^c -invariant Lagrangian $L_0 : TG \rightarrow \mathbb{R}$, the reduced Lagrangian l is defined on $\mathfrak{g} \times \mathcal{O}_0^c$ by

$$l(\xi, c(g^{-1})) = L_0(g\xi).$$

It will be sufficient to restrict to Lagrangians for simple mechanical systems with symmetry, that is, of the form $L_0(v_g) = K(v_g) - P(g)$, where K is the kinetic energy associated to a G_0^c -invariant Riemannian metric on G and the potential P is G_0^c -invariant. In this case, the reduced Lagrangian is

$$l(\xi, c(g^{-1})) = K(g\xi) - P(g).$$

Note that the right-hand side of this expression is well defined on $\mathfrak{g} \times \mathcal{O}_0^c$, that is, it depends on g only through $c(g^{-1})$. Indeed, $c(g^{-1}) = c(h^{-1})$ if and only if $\theta_{g^{-1}}(0) = \theta_{h^{-1}}(0)$, which means that $hg^{-1} \in G_0^c$. Therefore, $P(h) = P((hg^{-1})g) = P(g)$ by left G_0^c -invariance of P . For the kinetic energy the same argument works since the metric is G_0^c -invariant.

Thus we can write $L_0(v_g) = K(v_g) - E(c(g^{-1}))$ for the function $E : V^* \rightarrow \mathbb{R}$ uniquely determined by the relation $P(g) = E(c(g^{-1}))$. In this case, we have

$$l(\xi, c(g^{-1})) = K(g\xi) - E(c(g^{-1})).$$

For the Lagrangian of the charged molecular strand the potential energy is the sum of two terms, one of which, denoted by E_{loc} , explicitly depends only on $c(g^{-1})$ and the other, denoted by E_{np} , does not have a concrete expression only in terms of $c(g^{-1})$ but it is G_0^c -invariant. In addition, for the charged molecular strand the kinetic energy metric is not just G_0^c -invariant but G -invariant which then implies that it is only a function of $\xi \in \mathfrak{g}$. For the molecular strand the Lagrangian is of the form

$$L_0(v_g) = K(v_g) - E_{\text{loc}}(c(g^{-1})) - E_{\text{np}}(\zeta(g), c(g^{-1})),$$

where ζ is a G_0^c -invariant function defined on G and the reduced Lagrangian is

$$\begin{aligned} l(\xi, c(g^{-1})) &= \underbrace{K(\xi) - E_{\text{loc}}(c(g^{-1}))}_{=l_{\text{loc}}} - E_{\text{np}}(\zeta(g), c(g^{-1})) \\ &= l_{\text{loc}}(\xi, c(g^{-1})) + l_{\text{np}}(\zeta(g), c(g^{-1})). \end{aligned}$$

Note that l can be expressed in terms of $(\xi, a) \in \mathfrak{g} \times \mathcal{O}_0^c$ as

$$l(\xi, a) = K(\xi) - E_{\text{loc}}(a) - E_{\text{np}}(\zeta(g_a), a) = l_{\text{loc}}(\xi, a) + l_{\text{np}}(\zeta(g_a), a), \quad (5.16)$$

where $g_a \in G$ is such that $c(g_a^{-1}) = a$. This g_a is determined only up to left multiplication by G_0^c . Since E_{np} is G_0^c -invariant, the function $a \mapsto E_{\text{np}}(g_a)$ is well-defined. Because $a \mapsto l_{\text{np}}(\zeta(g_a), a)$ is a well-defined function of $a \in \mathcal{O}_0^c$ one can ask why we insist in denoting $l_{\text{np}} = l_{\text{np}}(\zeta(g_a), a)$ instead of simply $l_{\text{np}} = l_{\text{np}}(a)$ which is mathematically correct. The reason is that for the molecular strand we do not have an explicit expression for $l_{\text{np}} : \mathcal{O}_0^c \rightarrow \mathbb{R}$; see (2.14).

Let us identify all relevant objects that appear in the dynamics of the molecular strand as an example of this abstract setup. The Lie group is $G = \mathcal{F}([0, L], SE(3))$, its Lie algebra is $\mathfrak{g} = \mathcal{F}([0, L], \mathfrak{se}(3))$, the representation space is $V = \mathfrak{X}([0, L], \mathbb{R}^3) \oplus \mathcal{F}([0, L], \mathbb{R}^3)$, and its dual is $V^* = \Omega^1([0, L], \mathfrak{se}(3)) \oplus \mathcal{F}([0, L], \mathbb{R}^3)$.

The spaces $\Omega^1([0, L], \mathfrak{se}(3))$ and $\mathfrak{X}([0, L], \mathbb{R}^3)$ are vector valued one-forms, respectively vector fields (contravariant one-tensors). Of course, since $[0, L]$ is one-dimensional, both of these spaces are naturally identified with smooth functions with values in the respective vector spaces. But, in view of the generalization presented later in Section 5.4, it is useful to think of these spaces already in this fashion. Also, the various \mathbb{R}^3 appearing above play different roles: they can be the Lie algebra of $\mathfrak{so}(3)$, its dual, the natural representation space of $SO(3)$, or its dual. When discussing the generalization in Section 5.4, of course all these spaces are different.

The variables associated to these spaces are the following. Elements of G are denoted by (Λ, \mathbf{r}) , where $\Lambda \in \mathcal{F}([0, L], SO(3))$ and $\mathbf{r} \in \mathcal{F}([0, L], \mathbb{R}^3)$. Elements of \mathfrak{g} are denoted by $(\omega, \boldsymbol{\gamma})$, with $\omega, \boldsymbol{\gamma} \in \mathcal{F}([0, L], \mathbb{R}^3)$. Finally, elements of V^*

are denoted by $(\mathbf{\Omega}, \mathbf{\Gamma}, \rho)$, where $\mathbf{\Omega} \in \Omega^1([0, L], \mathbb{R}^3)$, $\mathbf{\Gamma} \in \Omega^1([0, L], \mathbb{R}^3)$, and $\rho \in C^\infty([0, L], \mathbb{R}^3)$.

The V^* -valued one-cocycle on G is given by

$$\begin{aligned} c((\Lambda, \mathbf{r})^{-1}) &= \left((\Lambda, \mathbf{r})^{-1} (\Lambda, \mathbf{r})', \Lambda^{-1} \mathbf{r} \right) \\ &= (\Lambda^{-1} \Lambda', \Lambda^{-1} \mathbf{r}', \Lambda^{-1} \mathbf{r}) = (\mathbf{\Omega}, \mathbf{\Gamma}, \rho). \end{aligned} \quad (5.17)$$

The function ζ appearing in l_{np} is given in this case by

$$\zeta(s, s') = (\xi(s, s'), \kappa(s, s')) = (\Lambda, \mathbf{r})^{-1}(s)(\Lambda, \mathbf{r})(s') \in SE(3),$$

so $\zeta(s, _) \in G$. Note that the Lagrangian of the strand (see (2.4), (2.13), and (2.14)) is exactly of the form (5.16), with $\zeta(s, s') = (\xi(s, s'), \kappa(s, s')) \in SE(3)$ given above. In fact, (2.14) has an expression of the type $l_{\text{np}} = l_{\text{np}}(\zeta(g_a), a)$. These comments will be greatly expanded and explained in detail in Section 5.4.

5.3.2. Recovering the modified Euler–Poincaré approach By Theorem 5.3, we have seen that the Euler–Lagrange equations of a G_0^c -invariant Lagrangian $L_0 : TG \rightarrow \mathbb{R}$ are equivalent to the affine Euler–Poincaré equations for $l : \mathfrak{g} \times \mathcal{O}_0^c \rightarrow \mathbb{R}$, that is,

$$\frac{\partial}{\partial t} \frac{\delta l}{\delta \xi} = \text{ad}_\xi^* \frac{\delta l}{\delta \xi} + \frac{\delta l}{\delta a} \diamond a - \mathbf{d}c^T \left(\frac{\delta l}{\delta a} \right). \quad (5.18)$$

Recall that to write these equations, we need to extend l to $\mathfrak{g} \times V^*$. Nevertheless, as we have shown, this extension does not affect the solution of these equations. For the molecular strand, there is an additional complication coming from the fact that the Lagrangian

$$l(\xi, a) = l_{\text{loc}}(\xi, a) + l_{\text{np}}(\zeta(g_a), a) \quad (5.19)$$

being a well defined function of $(\xi, a) \in \mathfrak{g} \times \mathcal{O}_0^c$, is not explicitly written in terms of a . Therefore, when computing the affine Euler–Poincaré equations in concrete examples, there is still a dependence on g_a in the final equation, although we know that this dependence can be replaced by a dependence in a uniquely, by the results above.

Let us apply the variational principle (5.14) to the Lagrangian in (5.19). Let $g(t)$ be a given curve in G . Take a family of curves $g_\varepsilon(t)$ satisfying $g_0(t) = g(t)$ and denote $\eta(t) := g^{-1}(t)\delta g(t) \in \mathfrak{g}$. Then $\delta \int_{t_0}^{t_1} l(\xi(t), c(g(t)^{-1}))dt = 0$ implies

$$\frac{\partial}{\partial t} \frac{\delta l_{\text{loc}}}{\delta \xi} = \text{ad}_\xi^* \frac{\delta l_{\text{loc}}}{\delta \xi} + \frac{\delta(l_{\text{loc}} + l_{\text{np}})}{\delta a} \diamond a - \mathbf{d}c^T \left(\frac{\delta(l_{\text{loc}} + l_{\text{np}})}{\delta a} \right) + g^{-1} \frac{\delta l_{\text{np}}}{\delta \zeta} T_g \zeta. \quad (5.20)$$

Note that this equation is the abstract generalization of equations (3.23) and (3.26).

Recall from the abstract theory that l_{np} depends only on $a \in \mathcal{O}_0^c$. However, l_{np} is given as a function of $(\zeta(g), c(g^{-1}))$. Let

$$\left. \frac{\delta l_{\text{np}}}{\delta a} \right|_{\text{Tot}}$$

denote the functional derivative of l_{np} viewed as a function of $a \in \mathcal{O}_0^c$ only. Since every curve in \mathcal{O}_0^c through $a = c(g^{-1}) \in \mathcal{O}_0^c$ is of the form $c(g_\varepsilon^{-1})$, where $g_0 = g$, we have

$$\begin{aligned} \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} l_{\text{np}}(\zeta(g_\varepsilon), c(g_\varepsilon^{-1})) &= \left\langle \left. \frac{\delta l_{\text{np}}}{\delta a} \right|_{\text{Tot}}, \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} c(g_\varepsilon^{-1}) \right\rangle \\ &= - \left\langle \left. \frac{\delta l_{\text{np}}}{\delta a} \right|_{\text{Tot}}, \eta a + \mathbf{d}c(\eta) \right\rangle \\ &= \left\langle \left. \frac{\delta l_{\text{np}}}{\delta a} \right|_{\text{Tot}}, \diamond a - \mathbf{d}c^T \left(\left. \frac{\delta l_{\text{np}}}{\delta a} \right|_{\text{Tot}} \right), \eta \right\rangle, \quad (5.21) \end{aligned}$$

where $\eta := g^{-1} \delta g$. On the other hand,

$$\begin{aligned} \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} l_{\text{np}}(\zeta(g_\varepsilon), c(g_\varepsilon^{-1})) &= \left\langle \frac{\delta l_{\text{np}}}{\delta \zeta}, T_g \zeta(g\eta) \right\rangle - \left\langle \frac{\delta l_{\text{np}}}{\delta a}, \eta a + \mathbf{d}c(\eta) \right\rangle \\ &= \left\langle g^{-1} \frac{\delta l_{\text{np}}}{\delta \zeta} T_g \zeta + \frac{\delta l_{\text{np}}}{\delta a} \diamond a - \mathbf{d}c^T \left(\frac{\delta l_{\text{np}}}{\delta a} \right), \eta \right\rangle. \quad (5.22) \end{aligned}$$

Equations (5.21) and (5.21) prove the following identity

$$\left. \frac{\delta l_{\text{np}}}{\delta a} \right|_{\text{Tot}} \diamond a - \mathbf{d}c^T \left(\left. \frac{\delta l_{\text{np}}}{\delta a} \right|_{\text{Tot}} \right) = \frac{\delta l_{\text{np}}}{\delta a} \diamond a - \mathbf{d}c^T \left(\frac{\delta l_{\text{np}}}{\delta a} \right) + g^{-1} \frac{\delta l_{\text{np}}}{\delta \zeta} T_g \zeta,$$

where $a = c(g^{-1})$. Using this identity in (5.20) we obtain the affine Euler–Poincaré equations (5.18) since

$$\frac{\delta l}{\delta a} = \left. \frac{\delta l_{\text{np}}}{\delta a} \right|_{\text{Tot}} + \frac{\delta l_{\text{loc}}}{\delta a}.$$

Thus, the affine Euler–Poincaré process recovers the results of the modified Euler–Poincaré approach described in Section 3.1.

5.3.3. Hamiltonian approach We now explore the Hamiltonian counterpart of the theory, that is, the case of a $G_{a_{\text{ref}}}^c$ -invariant Hamiltonian $H_{a_{\text{ref}}} : T^*G \rightarrow \mathbb{R}$, defined only for a fixed value $a_{\text{ref}} \in V^*$. As before, we do not suppose that $H_{a_{\text{ref}}}$ is induced from a G -invariant Hamiltonian on $T^*G \times V^*$. In particular, we do not know the expression of H_a for other choices of a . Such an $H_{a_{\text{ref}}}$ is usually induced by a hyperregular $G_{a_{\text{ref}}}^c$ -invariant Lagrangian $L_{a_{\text{ref}}}$.

As on the Lagrangian side, the reduced Hamiltonian is defined only on the submanifold

$$\mathfrak{g}^* \times \mathcal{O}_{a_{\text{ref}}}^c \subset \mathfrak{s}^*$$

and so Theorem 5.2 cannot be applied. However, as is shown in the next theorem, the fact that the reduced motion is Hamiltonian on an affine coadjoint orbit remains true for this more general case.

We need to introduce the affine coadjoint orbit $\mathcal{O}_{(\mu,a)}^\sigma$. The left V^* -valued group one-cocycle $c : G \rightarrow V^*$ induces a left group one-cocycle $\sigma : S \rightarrow (\mathfrak{g} \ltimes V)^*$ by

$$\sigma(g, u) = (u \diamond c(g) - \mathbf{d}c^T(u), c(g)).$$

The affine coadjoint action of S on \mathfrak{s}^* is hence given by

$$(g, u)(\mu, a) := \text{Ad}_{(g,u)^{-1}}^*(\mu, a) + \sigma((g, u)^{-1}),$$

where $g \in G, u \in V, \mu \in \mathfrak{g}^*$, and $a \in V^*$. The connected components of the coadjoint orbits $(\mathcal{O}_{(\mu,a_{\text{ref}})}^\sigma, \omega^-)$ are the symplectic leaves of \mathfrak{s}^* endowed with the affine Lie–Poisson bracket (5.10). Denote by $S_{(\mu,a)}^\sigma$ the isotropy group of the affine coadjoint action.

Theorem 5.4. *Let $H_{a_{\text{ref}}} : T^*G \rightarrow \mathbb{R}$ be a $G_{a_{\text{ref}}}^c$ -invariant Hamiltonian, where a_{ref} is a fixed element in V^* . By $G_{a_{\text{ref}}}^c$ -invariance, we obtain the reduced Hamiltonian h on $\mathfrak{g}^* \times \mathcal{O}_{a_{\text{ref}}}^c \subset \mathfrak{s}^*$ defined by $h(\mu, \theta_g(a_{\text{ref}})) = H_{a_{\text{ref}}}(g^{-1}\mu)$.*

- (i) *Let $\alpha(t) \in T_{g(t)}^*G$ be a solution of Hamilton's equations associated to $H_{a_{\text{ref}}}$ with initial condition $\mu_0 \in T_e^*G = \mathfrak{g}^*$. Then $(\mu(t), a(t)) := (g(t)^{-1}\alpha(t), \theta_{g(t)^{-1}}(a_{\text{ref}})) \in \mathfrak{s}^*$ is the integral curve of the Hamiltonian vector field X_h on the affine coadjoint orbit $(\mathcal{O}_{(\mu_0,a_{\text{ref}})}^\sigma, \omega^-)$ with initial condition (μ_0, a_0) . Conversely, given $\mu_0 \in \mathfrak{g}^* = T_e^*G$, the solution $\alpha(t)$ of the Hamiltonian system associated to $H_{a_{\text{ref}}}$ is reconstructed from the solution $(\mu(t), a(t))$ of $X_h \in \mathfrak{X}(\mathcal{O}_{(\mu_0,a_{\text{ref}})}^\sigma)$ with initial condition (μ_0, a_0) by setting $\alpha(t) = g(t)\mu(t)$, where $g(t)$ is the unique solution of the differential equation $\dot{g}(t) = g(t) \frac{\delta h}{\delta \mu(t)}$ with initial condition $g(0) = e$.*
- (ii) *Extending h arbitrarily to \mathfrak{s}^* , Hamilton's equations on $(\mathcal{O}_{(\mu_0,a_{\text{ref}})}^\sigma, \omega^-)$ can be written as*

$$\frac{\partial}{\partial t}(\mu, a) = \left(\text{ad}_{\frac{\delta h}{\delta \mu}}^* \mu - \frac{\delta h}{\delta a} \diamond a + \mathbf{d}c^T \left(\frac{\delta h}{\delta a} \right), -\frac{\delta h}{\delta \mu} a - \mathbf{d}c \left(\frac{\delta h}{\delta \mu} \right) \right)$$

where at the initial time $t = 0$, $\mu(0) = \mu_0$ and $a(0) = g(0)^{-1}a_{\text{ref}} + c(g(0)^{-1})$.

Remark 5.3. It is important to observe that the given Hamiltonian h is not defined on the whole dual \mathfrak{s}^* of the Lie algebra \mathfrak{s} . Part ii of the theorem states that the equations of motion can be nevertheless computed from the usual formula of an affine Lie–Poisson vector field by arbitrarily extending h to \mathfrak{s}^* . Note that $\delta h / \delta \mu \in \mathfrak{g}$ and $\delta h / \delta a \in V$ make sense only if one thinks of h as being defined on \mathfrak{s}^* . It will be shown in the proof of the theorem that the extension of h does not matter. This difficulty will appear concretely when dealing with the molecular strand.

Proof. (i) The action Ψ of S on T^*S induces an action of V on T^*S given by

$$(\alpha_h, (u, a)) \mapsto (\alpha_h, v + u, a),$$

where $h \in G$, $\alpha_h \in T^*G$, $u \in V$, and $a \in V^*$. Since V is a closed subgroup of S , this action admits a momentum map given by

$$\mathbf{J}_V(\alpha_g, (u, a)) = a.$$

Since V is an Abelian group, the coadjoint isotropy group of $a_{\text{ref}} \in V^*$ is $V_{a_{\text{ref}}} = V$ and the first reduced space $(T^*S)_{a_{\text{ref}}} = \mathbf{J}_V^{-1}(a_{\text{ref}})/V$ is symplectically diffeomorphic to the canonical symplectic manifold $(T^*G, \Omega_{\text{can}})$. The action Ψ of S on T^*S restricts to an action $\Psi^{a_{\text{ref}}}$ of $G_{a_{\text{ref}}}^c \curvearrowright V$ on $\mathbf{J}_V^{-1}(a_{\text{ref}})$. Passing to quotient spaces, this action induces an action of $G_{a_{\text{ref}}}^c$ on $(T^*S)_{a_{\text{ref}}}$, which is readily seen to be the cotangent lifted action of $G_{a_{\text{ref}}}^c$ on T^*G . We denote by $\mathbf{J}_{a_{\text{ref}}} : (T^*S)_{a_{\text{ref}}} \rightarrow (\mathfrak{g}_{a_{\text{ref}}}^c)^*$ the associated equivariant momentum map, where $\mathfrak{g}_{a_{\text{ref}}}^c$ is the Lie algebra of $G_{a_{\text{ref}}}^c$. Reducing $(T^*S)_{a_{\text{ref}}}$ at the point $\mu_{a_{\text{ref}}} := \mu|_{\mathfrak{g}_{a_{\text{ref}}}^c}$, we get the second reduced space $((T^*S)_{a_{\text{ref}}})_{\mu_{a_{\text{ref}}}} = \mathbf{J}_{a_{\text{ref}}}^{-1}(\mu_{a_{\text{ref}}})/(G_{a_{\text{ref}}}^c)_{\mu_{a_{\text{ref}}}}$, whose reduced symplectic form is denoted by $(\Omega_{a_{\text{ref}}})_{\mu_{a_{\text{ref}}}}$.

By the Reduction by Stages Theorem for nonequivariant momentum maps [28], the second reduced space is symplectically diffeomorphic to the reduced space

$$\left(\mathbf{J}^{-1}(\mu, a_{\text{ref}})/S_{(\mu, a_{\text{ref}})}^\sigma, \Omega_{(\mu, a_{\text{ref}})} \right)$$

obtained by reducing T^*S by the whole group S at the point $(\mu, a_{\text{ref}}) \in \mathfrak{s}^*$. By affine Lie–Poisson reduction, this space is symplectically diffeomorphic to the affine coadjoint orbit

$$\left(\mathcal{O}_{(\mu, a_{\text{ref}})}^\sigma, \omega^- \right)$$

endowed with the affine orbit symplectic form.

Note, finally, that by the symplectic reduction theorem, any solution of Hamilton's equations associated to $H_{a_{\text{ref}}}$ on T^*G reduces to and is reconstructed from a solution of Hamilton's equations for the reduced Hamiltonian $h_{\mu_{a_{\text{ref}}}} : \mathbf{J}_{a_{\text{ref}}}^{-1}(\mu_{a_{\text{ref}}})/(G_{a_{\text{ref}}}^c)_{\mu_{a_{\text{ref}}}} \rightarrow \mathbb{R}$, for a given momentum value $\mu_{a_{\text{ref}}} \in (\mathfrak{g}_{a_{\text{ref}}}^c)^*$. As we have seen, this reduced space is symplectically diffeomorphic to the affine coadjoint orbit $\mathcal{O}_{(\mu, a_{\text{ref}})}^\sigma \subset \mathfrak{s}^*$, where $\mu \in \mathfrak{g}^*$ is such that $\mu|_{\mathfrak{g}_{a_{\text{ref}}}^c} = \mu_{a_{\text{ref}}}$. Thus, we can think of $h_{\mu_{a_{\text{ref}}}}$ as being defined on $\mathcal{O}_{(\mu, a_{\text{ref}})}^\sigma$. Viewed this way, $h_{\mu_{a_{\text{ref}}}}$ is simply the restriction of the function h constructed from $H_{a_{\text{ref}}}$ by

$$h(\mu, \theta_g(a_{\text{ref}})) = H(g^{-1}\mu, a_{\text{ref}}).$$

Note that h is defined on any affine coadjoint orbit $\mathcal{O}_{(\mu, a_{\text{ref}})}^\sigma$ with fixed $a_0 \in V^*$ since

$$\mathfrak{g}^* \times \mathcal{O}_{a_{\text{ref}}}^c = \bigcup_{\mu \in \mathfrak{g}^*} \mathcal{O}_{(\mu, a_{\text{ref}})}^\sigma \subset \mathfrak{s}^*.$$

- (ii) We begin by recalling a general fact from the theory of Poisson manifolds. Let $\varphi \in C^\infty(P)$, where P is a Poisson manifold and X_φ its Hamiltonian vector field.

If L is a symplectic leaf of P , then $X_\varphi|_L = X_{(\varphi|_L)}$, where the left-hand side is the Hamiltonian vector field of φ on the manifold P , restricted to L , whereas the right-hand side denotes the Hamiltonian vector field on the symplectic manifold L relative to the Hamiltonian $\varphi|_L$ on L . In our case $P = \mathfrak{s}^*$ and $L = \mathcal{O}_{(\mu, a_{\text{ref}})}^\sigma$. \square

Remark 5.4. (The case $a_{\text{ref}} = 0$ and the molecular strand) The Lagrangian

$$L_0(v_g) = K(v_g) - E_{\text{loc}}(c(g^{-1})) - E_{\text{np}}(\zeta(g), c(g^{-1}))$$

discussed in Remark 5.2 is hyperregular, thus it induces the G_0^c -invariant Hamiltonian

$$H_0(\alpha_g) = K(\alpha_g) + E_{\text{loc}}(c(g^{-1})) + E_{\text{np}}(\zeta(g), c(g^{-1}))$$

whose reduced expression on $\mathfrak{g}^* \times \mathcal{O}_0^c$ reads

$$h(\mu, c(g^{-1})) = \frac{1}{2} \|\mu\|^2 + E_{\text{loc}}(c(g^{-1})) + E_{\text{np}}(\zeta(g), c(g^{-1})).$$

As on the Lagrangian side, for $(\mu, a) \in \mathfrak{g}^* \times \mathcal{O}_0^c$ (or $(\mu, a) \in \mathcal{O}_{(\mu_0, 0)}^\sigma$), we can write

$$h(\mu, a) = \frac{1}{2} \|\mu\|^2 + E_{\text{loc}}(a) + E_{\text{np}}(\zeta(g_a), a),$$

where $g_a \in G$ is any group element satisfying $c(g_a^{-1}) = a$.

Remark 5.5. (Affine coadjoint orbits and Noether’s theorem) As we have already seen, the solution (μ, a) evolves on an affine coadjoint orbit, for any $G_{a_{\text{ref}}}^c$ -invariant Hamiltonian $H_{a_{\text{ref}}}$. If $L_{a_{\text{ref}}}$ is the Lagrangian of a simple mechanical system with symmetry then, by Noether’s theorem, the solution (ξ, a) is constrained to evolve on the submanifolds

$$\left(\mathcal{O}_{(\mu_0, a_{\text{ref}})}^\sigma\right)^\sharp = \left\{(\xi, a) \in \mathfrak{g} \times V^* \mid (\xi^\flat, \mu) \in \mathcal{O}_{(\mu_0, a_{\text{ref}})}^\sigma\right\},$$

where $\flat : \mathfrak{g} \rightarrow \mathfrak{g}^*$ is defined by the kinetic energy metric and $\sharp : \mathfrak{g}^* \rightarrow \mathfrak{g}$ is its inverse.

5.4. Application to the charged strand

In this subsection we apply the affine Euler–Poincaré and Lie–Poisson reduction theorems to the charged strand.

Recall that the variables needed for this problem consist of the Lagrangian quantities $(\Lambda, \mathbf{r}) : [0, L] \rightarrow SE(3)$, together with the convective variables defined by

$$\Omega := \Lambda^{-1} \Lambda', \quad \omega := \Lambda^{-1} \dot{\Lambda} : [0, L] \rightarrow \mathfrak{so}(3)$$

and

$$\mathbf{\Gamma} := \Lambda^{-1} \mathbf{r}', \quad \mathbf{y} := \Lambda^{-1} \dot{\mathbf{r}}, \quad \boldsymbol{\rho} := \Lambda^{-1} \mathbf{r} : [0, L] \rightarrow \mathbb{R}^3.$$

In order to give a more transparent vision of the underlying geometric structures, we consider the n -dimensional generalization described in Section 2.4, that is, we replace the interval $[0, L]$ by an arbitrary manifold \mathcal{D} and we replace $SE(3)$ by the semidirect product $S = \mathcal{O} \ltimes E$ of a Lie group \mathcal{O} with a *left* representation space E . Given a manifold \mathcal{D} , we define the group $G := \mathcal{F}(\mathcal{D}, S)$ of smooth S -valued functions on \mathcal{D} and the dual vector space $V^* := \Omega^1(\mathcal{D}, \mathfrak{s}) \oplus \mathcal{F}(\mathcal{D}, E)$ consisting of pairs formed by smooth \mathfrak{s} -valued one-forms and E -valued functions on \mathcal{D} . The elements of the group G are denoted by (Λ, r) , where $\Lambda : \mathcal{D} \rightarrow \mathcal{O}$ and $r : \mathcal{D} \rightarrow E$. The elements of V^* are denoted by (Ω, Γ, ρ) , where $\Omega \in \Omega^1(\mathcal{D}, \mathfrak{o})$, $\Gamma \in \Omega^1(\mathcal{D}, E)$, and $\rho : \mathcal{D} \rightarrow E$. The space V^* can be seen as the dual of $V = \mathfrak{X}(\mathcal{D}, \mathfrak{s}^*) \oplus \mathcal{F}(\mathcal{D}, E^*)$, where $\mathfrak{X}(\mathcal{D}, \mathfrak{s})$ is the space of \mathfrak{s} -valued vector fields on \mathcal{D} .

Consider the representation of G on V^* defined by

$$(\Lambda, r)(\Omega, \Gamma, \rho) = (\text{Ad}_{(\Lambda, r)}(\Omega, \Gamma), \Lambda\rho) \quad (5.23)$$

where the adjoint action is that of S , acting here on functions defined on \mathcal{D} , and $\Lambda\rho$ denotes the left representation of \mathcal{O} on E , acting on functions. The main object for this approach is the *group one-cocycle* c appearing already implicitly in the definition of the variables $\boldsymbol{\Omega}, \mathbf{\Gamma}, \boldsymbol{\rho}$ in (2.11), (2.20), and explicitly in (2.21). Recall that it is given by (5.17), now rewritten as

$$c(\Lambda, r) := \left((\Lambda, r) \mathbf{d}(\Lambda, r)^{-1}, -r \right).$$

Let's verify the cocycle identity for the first component $(\Lambda, r) \mathbf{d}(\Lambda, r)^{-1}$. To simplify notation, denote $\chi_i := (\Lambda_i, r_i) \in \mathcal{F}(\mathcal{D}, S)$, $i = 1, 2$. We have

$$\begin{aligned} \chi_1 \chi_2 \mathbf{d}(\chi_1 \chi_2)^{-1} &= \chi_1 \chi_2 \mathbf{d}(\chi_2^{-1} \chi_1^{-1}) = \chi_1 \chi_2 \mathbf{d}(\chi_2^{-1}) \chi_1^{-1} + \chi_1 \chi_2 \chi_2^{-1} \mathbf{d}(\chi_1^{-1}) \\ &= \text{Ad}_{\chi_1} \left(\chi_2 \mathbf{d} \chi_2^{-1} \right) + \chi_1 \mathbf{d}(\chi_1^{-1}). \end{aligned}$$

Since the second coordinate of $((\Lambda_1, r_1)(\Lambda_2, r_2))$ is equal to $r = r_1 + \Lambda_1 r_2$, we find

$$\begin{aligned} c((\Lambda_1, r_1)(\Lambda_2, r_2)) &= \left(\text{Ad}_{(\Lambda_1, r_1)} \left((\Lambda_2, r_2) \mathbf{d}(\Lambda_2, r_2)^{-1} \right) + (\Lambda_1, r_1) \mathbf{d}((\Lambda_1, r_1)^{-1}), -r_1 - \Lambda_1 r_2 \right) \\ &= \left(\text{Ad}_{(\Lambda_1, r_1)} \left((\Lambda_2, r_2) \mathbf{d}(\Lambda_2, r_2)^{-1} \right), -\Lambda_1 r_2 \right) + \left((\Lambda_1, r_1) \mathbf{d}((\Lambda_1, r_1)^{-1}), -r_1 \right) \\ &= (\Lambda_1, r_1) c(\Lambda_2, r_2) + c(\Lambda_1, r_1). \end{aligned}$$

This shows that c verifies the cocycle property (5.2) relative to the representation (5.23).

Note that the first component of c is the left version of the cocycle appearing in the theory of complex fluids; see [12]. Using the expressions

$$\begin{aligned}(u, w, f) \diamond (\Omega, \Gamma, \rho) &= (\text{ad}_{\Omega_i}^* u^i + w^i \diamond \Gamma_i + f \diamond \rho, -\Omega_i w^i), \\ \mathbf{d}c(\omega, \gamma) &= (-\mathbf{d}\omega, -\mathbf{d}\gamma, -\gamma), \quad \text{and} \\ \mathbf{d}c^T(u, w, f) &= (\text{div}(u), \text{div}(w) - f),\end{aligned}$$

the *affine Euler–Poincaré equations* (5.8) become

$$\begin{cases} (\partial_t - \text{ad}_{\omega}^*) \frac{\delta l}{\delta \omega} + (\text{div} - \text{ad}_{\Omega}^*) \frac{\delta l}{\delta \Omega} = \frac{\delta l}{\delta \gamma} \diamond \gamma + \frac{\delta l}{\delta \Gamma} \diamond \Gamma + \frac{\delta l}{\delta \rho} \diamond \rho \\ (\partial_t + \omega) \frac{\delta l}{\delta \gamma} + (\text{div} + \Omega) \frac{\delta l}{\delta \Gamma} = \frac{\delta l}{\delta \rho} \end{cases} \quad (5.24)$$

and the *advection equations* are

$$\begin{cases} \partial_t \Omega + \text{ad}_{\omega} \Omega = \mathbf{d}\omega \\ (\partial_t + \omega) \Gamma = (\mathbf{d} + \Omega) \gamma \\ \partial_t \rho + \omega \rho = \gamma. \end{cases} \quad (5.25)$$

Remark 5.6. To write these equations, we have supposed that the dynamics is described by a Lagrangian l given explicitly in terms of the variables $(\omega, \gamma, \Omega, \Gamma, \rho)$. Equivalently, we have assumed that l is induced by an affine left-invariant Lagrangian L defined on $TG \times V^*$. As we have seen in Section 2.2, such a hypothesis is not verified when nonlocal terms are taken into account. In this case, the affine Euler–Poincaré and affine Lie–Poisson reductions are not applicable and one needs to restrict to a particular value of the parameter a_{ref} , by using Theorems 5.3 and 5.4. For convenience, we first present the simpler case where the nonlocal terms are ignored. We shall call this case *elastic filament dynamics*, for simplicity.

5.4.1. Elastic filament dynamics and Kirchhoff’s theory Suppose that the dynamics of the strand is described by a Lagrangian $l = l(\omega, \gamma, \Omega, \Gamma, \rho)$ defined on $\mathfrak{g} \times V^*$, where $\mathfrak{g} = \mathcal{F}(\mathcal{D}, \mathfrak{s})$ and $V^* = \Omega^1(\mathcal{D}, \mathfrak{s}) \oplus \mathcal{F}(\mathcal{D}, E)$. The Lagrangian l is induced by a left invariant Lagrangian L defined on $TG \times V^*$, where $G = \mathcal{F}(\mathcal{D}, S)$.

Note that there is no restriction in the way l depends on the variables. In particular the dependence can be nonlocal. However, it is supposed here that l depends explicitly on the variables $(\omega, \gamma, \Omega, \Gamma, \rho)$. Recall that such a hypothesis is verified for the Lagrangian of Kirchhoff’s theory (1.12) but is not verified for the Lagrangian of the charged strand (2.4).

The affine Euler–Poincaré reduction applies as follows. Fix the reference values $(\Omega_{\text{ref}}, \Gamma_{\text{ref}}, \rho_{\text{ref}})$ and define the Lagrangian

$$L_{(\Omega_{\text{ref}}, \Gamma_{\text{ref}}, \rho_{\text{ref}})}(\Lambda, r) := L(\Lambda, r, \Omega_{\text{ref}}, \Gamma_{\text{ref}}, \rho_{\text{ref}}).$$

Consider a curve $(\Lambda, r) \in G$ and define the quantities

$$\begin{aligned}(\Omega, \Gamma, \rho) &= (\Lambda, r)^{-1}(\Omega_{\text{ref}}, \gamma_{\text{ref}}, \rho_{\text{ref}}) + c((\Lambda, r)^{-1}) \\ &= (\text{Ad}_{\Lambda^{-1}} \Omega_{\text{ref}}, \Lambda^{-1}(\Gamma_{\text{ref}} + \Omega_{\text{ref}} r), \Lambda^{-1} \rho_{\text{ref}}) + (\Lambda^{-1} \mathbf{d}\Lambda, \Lambda^{-1} \mathbf{d}r, \Lambda^{-1} r).\end{aligned}$$

and

$$\omega = \Lambda^{-1} \dot{\Lambda}, \quad \gamma = \Lambda^{-1} \dot{\mathbf{r}}.$$

Note that when the initial values $\Omega_{\text{ref}}, \Gamma_{\text{ref}}, \rho_{\text{ref}}$ are zero, the definitions of the variables $\omega, \gamma, \Omega, \Gamma, \rho$ coincide with those given in (2.11) and (2.20).

Then the curve (Λ, \mathbf{r}) is a solution of the Euler–Lagrange equations associated to $L_{(\Omega_{\text{ref}}, \Gamma_{\text{ref}}, \rho_0)}$ on TG if and only if $(\omega, \gamma, \Omega, \Gamma, \rho)$ is a solution of the Euler–Poincaré equations (5.24).

Of course, when \mathcal{D} is the interval $[0, L]$ and S is the semidirect product of $\mathcal{O} = SO(3)$ with $E = \mathbb{R}^3$, then we recover from (5.24) the dynamical equation of the charged strand (3.38), since

$$\text{ad}^* \mapsto - \times \quad \text{and} \quad \diamond \mapsto \times.$$

These equations are the convective representation of Kirchhoff’s equations. From (5.25) we recover the advection relations derived in Section 2.3.

5.4.2. The charged strand: general case Recall from Section 2.2 that the Lagrangian of the charged strand has the expression

$$l = l_{\text{loc}}(\omega, \gamma, \Omega, \Gamma, \rho) + l_{\text{np}}(\xi, \kappa, \Gamma),$$

where l_{loc} is a local function of the form

$$l_{\text{loc}}(\omega, \gamma, \Omega, \Gamma, \rho) = K(\omega, \gamma) - E_{\text{loc}}(\Omega, \Gamma, \rho) \quad (5.26)$$

and l_{np} is of the form

$$l_{\text{np}}(\xi, \kappa, \Gamma) = \iint U(\xi(s, s'), \kappa(s, s'), \Gamma(s), \Gamma(s')) \, ds \, ds',$$

where

$$U : SE(3) \times \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R} \quad \text{and} \quad (\xi(s, s'), \kappa(s, s')) := (\Lambda, \mathbf{r})^{-1}(s)(\Lambda, \mathbf{r})(s').$$

Remark 5.7. (Two crucial observations)

1. The nonlocal Lagrangian l_{np} is induced by a $SO(3)$ -invariant potential $E_{\text{np}} = E_{\text{np}}(\Lambda, \mathbf{r})$. Thus the total Lagrangian l can be seen as being induced by the $SO(3)$ -invariant Lagrangian $L_0 = L_0(\Lambda, \dot{\Lambda}, \mathbf{r}, \dot{\mathbf{r}})$ given by

$$L_0(\Lambda, \dot{\Lambda}, \mathbf{r}, \dot{\mathbf{r}}) = K(\Lambda, \dot{\Lambda}, \mathbf{r}, \dot{\mathbf{r}}) - E_{\text{loc}}\left(c\left((\Lambda, \mathbf{r})^{-1}\right)\right) - E_{\text{np}}(\Lambda, \mathbf{r}),$$

where K is the $\mathcal{F}(\mathcal{D}, SE(3))$ -left invariant extension of the kinetic energy K in (5.26). Note that we have replaced the dependence of E_{loc} on (Ω, Γ, ρ) by a dependence on (Λ, \mathbf{r}) through the cocycle c . The affine Euler–Poincaré dynamics yields the relation $(\Omega, \Gamma, \rho) = c((\Lambda, \mathbf{r})^{-1})$ which allows us to recover the dependence of the potential on (Ω, Γ, ρ) .

2. The group $SO(3)$ is precisely the isotropy group

$$G_0^c = \mathcal{F}(\mathcal{D}, SE(3))_0^c = \{(\Lambda, r) \in G \mid c(\Lambda, r) = 0\}$$

of the affine action at zero.

These two remarks allow us to obtain the dynamics of the molecular strand by the affine reduction processes described in Theorems 5.3 and 5.4. As before, we choose to work with the general framework involving \mathcal{D} and $\mathcal{O} \otimes E$. The present approach is applicable to any \mathcal{O} -invariant Lagrangian

$$L_0 = L_0(\Lambda, \dot{\Lambda}, r, \dot{r}) : T[\mathcal{F}(\mathcal{D}, \mathcal{O} \otimes E)] \rightarrow \mathbb{R}.$$

Note there are no conditions on the dependence of L_0 on the variables (Λ, r) . In particular, L_0 can be nonlocal and may depend on the derivatives of Λ and r . An important class of such Lagrangians is given by

$$L_0(\Lambda, \dot{\Lambda}, r, \dot{r}) = K(\Lambda, \dot{\Lambda}, r, \dot{r}) - P(\Lambda, r),$$

where K is the kinetic energy associated to an \mathcal{O} -invariant metric on $\mathcal{F}(\mathcal{D}, \mathcal{O} \otimes E)$ and the potential P is an \mathcal{O} -invariant function on $\mathcal{F}(\mathcal{D}, \mathcal{O} \otimes E)$. In particular, P can be nonlocal, or can depend on derivatives of Λ and r ; see (1.12) for an example. In the case of the molecular strand, K is assumed to be left-invariant and P is given by

$$P(\Lambda, r) = E_{\text{loc}} \left(c \left((\Lambda, r)^{-1} \right) \right) + E_{\text{np}}(\Lambda, r),$$

where

$$E_{\text{np}}(\Lambda, r) := \iint_{\mathcal{D}} U \left(\xi(s, s'), \kappa(s, s'), \Lambda^{-1} \mathbf{d}r(s), \Lambda^{-1} \mathbf{d}r(s') \right) \mathrm{d}s \mathrm{d}s' \\ (\xi(s, s'), \kappa(s, s')) := (\Lambda, r)^{-1}(s)(\Lambda, r)(s') \in \mathcal{O} \otimes E$$

and one readily sees that E_{np} is \mathcal{O} -invariant. Recall that the cocycle is

$$c \left((\Lambda, r)^{-1} \right) = \left(\Lambda^{-1} \mathbf{d}\Lambda, \Lambda^{-1} \mathbf{d}r, \Lambda^{-1} r \right).$$

Thus, a straightforward and maybe useful generalization of E_{np} is

$$E_{\text{np}}(\Lambda, r) := \iint_{\mathcal{D}} U \left(\xi(s, s'), \kappa(s, s'), c \left((\Lambda, r)^{-1} \right) (s), c \left((\Lambda, r)^{-1} \right) (s') \right) \mathrm{d}s \mathrm{d}s'.$$

Using Theorem 5.3 with L_0 we obtain the same affine Euler–Poincaré equations (5.24), where all derivatives are total derivatives. One can equivalently use the modified Euler–Poincaré approach and obtain the equations

$$\left\{ \begin{aligned} (\partial_t - \text{ad}_\omega^*) \frac{\delta l}{\delta \omega} + (\text{div} - \text{ad}_\Omega^*) \frac{\delta l}{\delta \Omega} &= \frac{\delta l}{\delta \gamma} \diamond \gamma + \frac{\delta l}{\delta \Gamma} \diamond \Gamma + \frac{\delta l}{\delta \rho} \diamond \rho, \\ &+ \int \left[\xi(s, s') \frac{\partial U}{\partial \xi}(s', s) - \frac{\partial U}{\partial \xi}(s, s') \xi(s', s) - \kappa(s, s') \diamond \frac{\partial U}{\partial \kappa}(s, s') \right] \mathrm{d}s' \\ (\partial_t + \omega) \frac{\delta l}{\delta \gamma} + (\text{div} + \Omega) \frac{\delta l}{\delta \Gamma} &= \frac{\delta l}{\delta \rho} + \int \left[\xi(s, s') \frac{\partial U}{\partial \kappa}(s', s) - \frac{\partial U}{\partial \kappa}(s, s') \right] \mathrm{d}s'. \end{aligned} \right. \quad (5.27)$$

Note that here the derivatives are not total derivatives; see the discussion in Section 5.3.2. One can treat the Hamiltonian side in a similar way. As we have seen, the motion is Hamiltonian on affine coadjoint orbits.

5.4.3. Conservation laws and spatial formulation In this paragraph, we generalize the approach of Section 3.2 and reformulate the equations (5.24) for the generalized charged strand as a conservation law. We first need an n -dimensional generalization of formula (3.29).

Given a Lie group G , a map $g : \mathcal{D} \rightarrow G$ defined on a n -dimensional manifold \mathcal{D} , $s \in \mathcal{D}$, and a \mathfrak{g}^* -valued vector field w on \mathcal{D} , we have

$$\mathrm{Ad}_g^* \left[\mathrm{div} \left(\mathrm{Ad}_{g^{-1}}^* w \right) \right] = \mathrm{div} w - \mathrm{ad}_{\sigma_i}^* w^i =: \mathrm{div}^\sigma w, \quad \sigma := g^{-1} \mathbf{d}g \in \Omega^1(\mathcal{D}, \mathfrak{g}). \quad (5.28)$$

Using this formula, (3.29), the expression of ad^* associated to the semidirect product $\mathcal{O} \ltimes E$, and the equalities

$$(\omega, \gamma) = (\Lambda, r)^{-1}(\dot{\Lambda}, \dot{r}), \quad (\Omega, \Gamma) = (\Lambda, r)^{-1} \mathbf{d}(\Lambda, r),$$

we find

$$\begin{aligned} & \mathrm{Ad}_{(\Lambda, r)}^* \frac{\partial}{\partial t} \left[\mathrm{Ad}_{(\Lambda, r)^{-1}}^* \left(\frac{\delta l}{\delta \omega}, \frac{\delta l}{\delta \gamma} \right) \right] \\ &= \frac{\partial}{\partial t} \left(\frac{\delta l}{\delta \omega}, \frac{\delta l}{\delta \gamma} \right) + \left(-\mathrm{ad}_\omega^* \frac{\delta l}{\delta \gamma} + \gamma \diamond \frac{\delta l}{\delta \gamma}, \omega \frac{\delta l}{\delta \gamma} \right), \end{aligned}$$

and

$$\begin{aligned} & \mathrm{Ad}_{(\Lambda, r)}^* \mathrm{div} \left[\mathrm{Ad}_{(\Lambda, r)^{-1}}^* \left(\frac{\delta l}{\delta \Omega}, \frac{\delta l}{\delta \Gamma} \right) \right] \\ &= \mathrm{div} \left(\frac{\delta l}{\delta \Omega}, \frac{\delta l}{\delta \Gamma} \right) + \left(-\mathrm{ad}_\Omega^* \frac{\delta l}{\delta \Omega} + \Gamma \diamond \frac{\delta l}{\delta \Gamma}, \Omega \frac{\delta l}{\delta \Gamma} \right). \end{aligned}$$

Thus, equations (5.27) can be rewritten in the form of a conservation law, namely

$$\begin{aligned} & \frac{\partial}{\partial t} \left[\mathrm{Ad}_{(\Lambda, r)^{-1}}^* \left(\frac{\delta l}{\delta \omega}, \frac{\delta l}{\delta \gamma} \right) \right] + \mathrm{div} \left[\mathrm{Ad}_{(\Lambda, r)^{-1}}^* \left(\frac{\delta l}{\delta \Omega}, \frac{\delta l}{\delta \Gamma} \right) \right] \\ &= \mathrm{Ad}_{(\Lambda, r)^{-1}}^* \left(\frac{\delta l}{\delta \rho} \diamond \rho, \frac{\delta l}{\delta \rho} \right). \end{aligned} \quad (5.29)$$

Using (5.1), the right-hand side simplifies to

$$\begin{aligned} \mathrm{Ad}_{(\Lambda, r)^{-1}}^* \left(\frac{\delta l}{\delta \rho} \diamond \rho, \frac{\delta l}{\delta \rho} \right) &= \left(\mathrm{Ad}_{\Lambda^{-1}}^* \left(\frac{\delta l}{\delta \rho} \diamond \rho \right) + r \diamond \left(\Lambda \frac{\delta l}{\delta \rho} \right), \Lambda \frac{\delta l}{\delta \rho} \right) \\ &= \left(\left(\Lambda \frac{\delta l}{\delta \rho} \diamond \Lambda \rho \right) + r \diamond \left(\Lambda \frac{\delta l}{\delta \rho} \right), \Lambda \frac{\delta l}{\delta \rho} \right) = \left(0, \Lambda \frac{\delta l}{\delta \rho} \right), \end{aligned}$$

since $\rho = \Lambda^{-1}r$. Note that this is the exact equivalent of the simplification (1.24) derived at the beginning of the paper.

Such a conservation law is valid for each solution of the affine Euler–Poincaré equation (5.1) associated to a G_0^c -invariant Lagrangian $L_0 : TG \rightarrow \mathbb{R}$. In particular, it is valid for the Kirchhoff theory, as we saw at end of Section 1.3.1.

A short computation shows that, in general, the previous conservation law reads

$$\frac{\partial}{\partial t} \left[\text{Ad}_{g^{-1}}^* \frac{\delta l}{\delta \xi} \right] + \mathbf{d}c^T \left(g \frac{\delta l}{\delta a} \right) = 0. \quad (5.30)$$

When a_{ref} is not necessarily zero, the previous formula becomes

$$\frac{\partial}{\partial t} \left[\text{Ad}_{g^{-1}}^* \frac{\delta l}{\delta \xi} \right] + \mathbf{d}c^T \left(g \frac{\delta l}{\delta a} \right) = \text{Ad}_{g^{-1}}^* \left(\frac{\delta l}{\delta a} \diamond g^{-1} a_{\text{ref}} \right). \quad (5.31)$$

5.4.4. The fixed filament and its conservation law The equations (1.29) for a fixed filament can also be obtained by affine Euler–Poincaré reduction. It suffices to apply Theorem 5.3 with the group $G = \mathcal{F}(\mathcal{D}, \mathcal{O}) \ni \Lambda$, acting on the vector space $\Omega^1(\mathcal{D}, \mathfrak{o}) \times \mathcal{F}(\mathcal{D}, E) \ni (\Omega, \rho)$ by the affine action

$$(\Omega, \rho) \mapsto \theta_\Lambda(\Omega, \rho) := (\text{Ad}_\Lambda \Omega + \Lambda \mathbf{d}\Lambda^{-1}, \Lambda \rho).$$

Note that the cocycle is $c(\Lambda) = (\Lambda \mathbf{d}\Lambda^{-1}, 0)$. Using the expressions

$$(u, f) \diamond (\Omega, \rho) = \text{ad}_{\Omega_i}^* u^i + f \diamond \rho,$$

$$\mathbf{d}c(\omega) = (-\mathbf{d}\omega, 0), \quad \text{and} \quad \mathbf{d}c^T(u, f) = \text{div}(u),$$

the affine Euler–Poincaré equations (5.8) become

$$(\partial_t - \text{ad}_\omega^*) \frac{\delta l}{\delta \omega} + (\text{div} - \text{ad}_\Omega^*) \frac{\delta l}{\delta \Omega} = \frac{\delta l}{\delta \rho} \diamond \rho \quad (5.32)$$

and the advection equations are

$$\begin{cases} \partial_t \Omega + \text{ad}_\omega \Omega = \mathbf{d}\omega, \\ \partial_t \rho + \omega \rho = 0. \end{cases} \quad (5.33)$$

Recall from Section 1.3.2 that the Lagrangian for a fixed filament is of the form

$$l = l_{\text{loc}}(\omega, \Omega) + l_{\text{np}}(\xi, \rho),$$

$$l_{\text{loc}}(\omega, \Omega) = K(\omega) - \frac{1}{2} \int f(\Omega(s)) ds, \quad l_{\text{np}}(\xi, \rho) = - \iint U(\rho(s), \xi(s, s')) ds ds'$$

where

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}, \quad U : \mathbb{R}^3 \times SO(3) \rightarrow \mathbb{R}, \quad \xi(s, s') := \Lambda^{-1}(s) \Lambda(s').$$

Using the relations $\omega = \Lambda^{-1} \dot{\Lambda}$, $\Omega = \Lambda^{-1} \Lambda'$, and $\rho = \Lambda^{-1} \rho_{\text{ref}}$, where $\rho_{\text{ref}}(s) := \mathbf{r}(s) = (s, 0, 0)^T$, we conclude that l is induced by a $SO(2)$ -invariant Lagrangian $L_{(0, \mathbf{r})} = L_{(0, \mathbf{r})}(\Lambda, \dot{\Lambda})$. Note that $SO(2)$ is precisely the isotropy group of $(0, \mathbf{r})$ relative to the affine action.

These observations allow us to obtain the equations for the fixed filament by the affine reduction processes described in Theorems 5.3 and 5.4. Using the general framework involving \mathcal{D} and $\mathcal{O} \ltimes E$, we obtain the equations

$$\begin{aligned} & (\partial_t - \text{ad}_\omega^*) \frac{\delta l}{\delta \omega} + (\text{div} - \text{ad}_\Omega^*) \frac{\delta l}{\delta \Omega} \\ &= \frac{\delta l}{\delta \rho} \diamond \rho + \int \left[\xi(s, s') \frac{\partial U}{\partial \xi}(s', s) - \frac{\partial U}{\partial \xi}(s, s') \xi(s', s) \right] ds' \end{aligned}$$

which coincides with (1.28) in the case of the fixed filament. Using total derivatives, these equations can be rewritten as (5.32).

The general formula (5.31) yields the conservation law

$$\frac{\partial}{\partial t} \left[\text{Ad}_{\Lambda^{-1}}^* \frac{\delta l}{\delta \omega} \right] + \text{div} \left[\text{Ad}_{\Lambda^{-1}}^* \frac{\delta l}{\delta \Omega} \right] = \text{Ad}_{\Lambda^{-1}}^* \left(\frac{\delta l}{\delta \rho} \diamond \rho \right).$$

From the general theory, it follows that the solution of the advection equations (5.33) in terms of Λ are given by $\Omega = \Lambda^{-1} \mathbf{d}\Lambda$ and $\rho = \Lambda^{-1} \rho_{\text{ref}}$.

For the fixed filament, we choose $\mathcal{D} = [0, L]$, $E = \mathbb{R}^3$, $\mathcal{O} = SO(3)$, $\rho_{\text{ref}}(s) = \mathbf{r}(s) = (s, 0, 0)^T$ and we get

$$\frac{\partial}{\partial t} \left[\text{Ad}_{\Lambda^{-1}}^* \frac{\delta l}{\delta \omega} \right] + \frac{\partial}{\partial s} \left[\text{Ad}_{\Lambda^{-1}}^* \frac{\delta l}{\delta \Omega} \right] = \text{Ad}_{\Lambda^{-1}}^* \left(\frac{\delta l}{\delta \rho} \times \rho \right). \quad (5.34)$$

Note that in this case, the torque does not vanish. The explanation is that the reference value ρ_{ref} of ρ is not zero, so we need to use (5.31) instead of (5.30).

Observe that we can write

$$\text{Ad}_{\Lambda^{-1}}^* \left(\frac{\delta l}{\delta \rho} \times \rho \right) = \Lambda \frac{\delta l}{\delta \rho} \times \Lambda \rho = \Lambda \frac{\delta l}{\delta \rho} \times \begin{pmatrix} s \\ 0 \\ 0 \end{pmatrix}.$$

More generally, the right-hand side is

$$\left(\Lambda \frac{\delta l}{\delta \rho} \right) \times \mathbf{r},$$

where \mathbf{r} describes the fixed filament.

Note that the conservation law (5.34) does not appear in Section 1.3.2. It is a particular case of the general formula (5.31). We believe that the derivation of this law through the affine Euler–Poincaré theory is interesting and shows the breadth of application of our theories.

6. New variables: coordinate change and horizontal-vertical split

In this section, we show that a drastic simplification of the equations arises under a particular change of variables. We first consider the case of strands. This change of variables will then be extended to the general setting of the previous section where $[0, L]$ is replaced by a manifold \mathcal{D} and $SE(3)$ by an arbitrary semidirect product associated to a representation.

6.1. Motivation in terms of covariant derivatives

We can see from (1.1) that ρ , Γ , and γ satisfy the following relations

$$(\partial_s + \Omega \times) \rho = \Gamma, \quad (\partial_t + \omega \times) \rho = \gamma. \quad (6.1)$$

Thus the reduced variables (1.1) lead naturally to two differential operators

$$\frac{D}{Ds} = (\partial_s + \Omega \times), \quad \frac{D}{Dt} = (\partial_t + \omega \times), \quad (6.2)$$

which will be interpreted later as covariant derivatives, one with respect to space and the other with respect to time.

With this interpretation, we regard Γ and γ as covariant tangent vectors above ρ ,

$$\frac{D\rho}{Ds} = \Gamma, \quad \frac{D\rho}{Dt} = \gamma. \quad (6.3)$$

The operators from (6.2) also appear in the equations of motion (3.38), since we can write the second Euler–Poincaré equation in the form

$$\frac{D}{Dt} \frac{\delta l}{\delta \gamma} + \frac{D}{Ds} \frac{\delta l}{\delta \Gamma} - \frac{\delta l}{\delta \rho} = 0. \quad (6.4)$$

Taking (6.3) and (6.4) together, we see that (6.4) is in the form of the Euler–Lagrange equations where the partial derivatives have been replaced by covariant derivatives. With this interpretation in mind we can ask whether, by a change of variables, we can transform (6.4) to the canonical Euler–Lagrange form. In this section we find that such a change of variables does exist and we give it explicitly. This line of enquiry leads us to consider in the subsequent sections how the two sets of coordinates are related from a geometric point of view.

6.2. The case of charged strands

Rearranging equations (6.1), we find that the time and space derivatives of ρ are given by

$$\partial_s \rho = \Gamma - \Omega \times \rho, \quad \partial_t \rho = \gamma - \omega \times \rho.$$

Therefore, we introduce the coordinate change

$$\begin{aligned} \mathcal{F}(I, \mathfrak{so}(3)) \times \mathcal{F}(I, \mathbb{R}^3) \times \Omega^1(I, \mathfrak{so}(3)) \times \Omega^1(I, \mathbb{R}^3) \times \mathcal{F}(I, \mathbb{R}^3) \\ \ni (\omega, \gamma, \Omega, \Gamma, \rho) \mapsto (\rho, \rho_s, \rho_t, \omega, \Omega) \\ \in \mathcal{F}(I, \mathbb{R}^3) \times \Omega^1(I, \mathbb{R}^3) \times \mathcal{F}(I, \mathbb{R}^3) \times \mathcal{F}(I, \mathfrak{so}(3)) \times \Omega^1(I, \mathfrak{so}(3)), \end{aligned} \quad (6.5)$$

where we have defined two new variables

$$\rho_s := \Gamma - \Omega \times \rho, \quad \rho_t := \gamma - \omega \times \rho. \quad (6.6)$$

This candidate coordinate change allows us to transform the differential operators D/Ds and D/Dt into partial derivatives. We shall show that the equations of motion (3.23) and (3.26) have simple expressions if one uses this change of variables. As far as we know, this transformation has not been noticed before, in either nonlocal or local setting.

Notation 6.1. Denote by \bar{l} the Lagrangian l in terms of the new variables given by (6.5), that is,

$$l_{\text{loc}}(\omega, \gamma, \Omega, \Gamma, \rho) + l_{\text{np}}(\xi, \kappa, \Gamma) = \bar{l}_{\text{loc}}(\rho, \rho_s, \rho_t, \omega, \Omega) + \bar{l}_{\text{np}}(\xi, \kappa, \rho_s, \rho, \Omega).$$

6.3. Change of coordinates

The action principle for the local part of the Lagrangian \bar{l}_{loc} yields

$$\begin{aligned} 0 &= \delta \int l(\omega, \gamma, \Omega, \Gamma, \rho) dt \\ &= \int \left[\left\langle \frac{\delta l}{\delta \rho}, \delta \rho \right\rangle + \left\langle \frac{\delta l}{\delta \gamma}, \delta \gamma \right\rangle + \left\langle \frac{\delta l}{\delta \Gamma}, \delta \Gamma \right\rangle + \left\langle \frac{\delta l}{\delta \omega}, \delta \omega \right\rangle + \left\langle \frac{\delta l}{\delta \Omega}, \delta \Omega \right\rangle \right] dt \\ &= \delta \int \bar{l}(\rho, \rho_s, \rho_t, \omega, \Omega) dt \\ &= \int \left[\left\langle \frac{\delta \bar{l}}{\delta \rho}, \delta \rho \right\rangle + \left\langle \frac{\delta \bar{l}}{\delta \rho_s}, \delta \rho_s \right\rangle + \left\langle \frac{\delta \bar{l}}{\delta \rho_t}, \delta \rho_t \right\rangle + \left\langle \frac{\delta \bar{l}}{\delta \omega}, \delta \omega \right\rangle + \left\langle \frac{\delta \bar{l}}{\delta \Omega}, \delta \Omega \right\rangle \right] dt. \end{aligned} \quad (6.7)$$

Define free variations $\Psi(s) = \Lambda(s)^{-1} \delta r(s)$ and $\Sigma(s) = \Lambda(s)^{-1} \delta \Lambda(s)$. As usual, Ψ denotes the antisymmetric matrix that is obtained from Ψ by the hat map. Then, the following theorem holds.

Theorem 6.2. The variations in $\delta \rho_s$ and $\delta \rho_t$ yield dynamical equations for the local Lagrangian \bar{l}_{loc} in the following form

$$(\partial_s + \Omega \times) \frac{\delta \bar{l}_{\text{loc}}}{\delta \Omega} + (\partial_t + \omega \times) \frac{\delta \bar{l}_{\text{loc}}}{\delta \omega} = 0, \quad (6.8)$$

$$\frac{\delta \bar{l}_{\text{loc}}}{\delta \rho} - \partial_t \frac{\delta \bar{l}_{\text{loc}}}{\delta \rho_t} - \partial_s \frac{\delta \bar{l}_{\text{loc}}}{\delta \rho_s} = 0. \quad (6.9)$$

Remark 6.1. The derivatives in the equations (6.8) and (6.9) have now formally decoupled, although the equations themselves must be solved simultaneously because the Lagrangian l depends on all the variables. Also note that equation (6.9) is equivalent, for local Lagrangians, to (3.26) with the covariant derivatives replaced by partial derivatives (but relative to the new variables). This gives a new interpretation to the right-hand side of (3.23) as being terms that arise from the induced covariant derivative.

Proof. First, variations $\delta \rho_t$ and $\delta \rho_s$ are computed from (6.6) as follows:

$$\delta \rho_t = \delta \gamma - \delta \omega \times \rho - \omega \times \delta \rho, \quad (6.10)$$

$$\delta \rho_s = \delta \Gamma - \delta \Omega \times \rho - \Omega \times \delta \rho. \quad (6.11)$$

Then, using the identities

$$\delta \omega = \dot{\Sigma} + \omega \times \Sigma,$$

$$\delta \Omega = \Sigma' + \Omega \times \Sigma,$$

$$\delta \rho = -\Sigma \times \rho + \Psi,$$

$$\delta \gamma = \dot{\Psi} + \omega \times \Psi - \Sigma \times \gamma,$$

$$\delta \Gamma = \Psi' + \Omega \times \Psi - \Sigma \times \Gamma,$$

we find, for example, from the term involving the derivatives with respect to ρ_t ,

$$\begin{aligned} \left\langle \frac{\delta \bar{l}}{\delta \rho_t}, \delta \rho_t \right\rangle &= \left\langle \frac{\delta \bar{l}}{\delta \rho_t}, \dot{\Psi} + \omega \times \Psi - \Sigma \times \gamma \right. \\ &\quad \left. - (\dot{\Sigma} + \omega \times \Sigma) \times \rho - \omega \times (-\Sigma \times \rho + \Psi) \right\rangle \\ &= \left\langle -\frac{\partial}{\partial t} \frac{\delta \bar{l}}{\delta \rho_t}, \Psi \right\rangle \\ &\quad + \left\langle \frac{\partial}{\partial t} \left(\rho \times \frac{\delta \bar{l}}{\delta \rho_t} \right) - \gamma \times \frac{\delta \bar{l}}{\delta \rho_t} - (\rho \times \omega) \times \frac{\delta \bar{l}}{\delta \rho_t}, \Sigma \right\rangle, \quad (6.12) \end{aligned}$$

where we have used the Jacobi identity simplifying two triple cross products. We now employ the kinematic condition for the derivative of ρ ,

$$\partial_t \rho = \gamma - \omega \times \rho,$$

to simplify the Σ term in (6.12) and obtain the following simple condition

$$\left\langle \frac{\delta \bar{l}}{\delta \rho_t}, \delta \rho_t \right\rangle = \left\langle -\frac{\partial}{\partial t} \frac{\delta \bar{l}}{\delta \rho_t}, \Psi \right\rangle + \left\langle -\rho \times \frac{\partial}{\partial t} \frac{\delta \bar{l}}{\delta \rho_t}, \Sigma \right\rangle. \quad (6.13)$$

Analogously,

$$\left\langle \frac{\delta \bar{l}}{\delta \rho_s}, \delta \rho_s \right\rangle = \left\langle -\frac{\partial}{\partial s} \frac{\delta \bar{l}}{\delta \rho_s}, \Psi \right\rangle + \left\langle -\rho \times \frac{\partial}{\partial s} \frac{\delta \bar{l}}{\delta \rho_s}, \Sigma \right\rangle. \quad (6.14)$$

On completing the variational principle (6.7) for all variables, one sees that the only terms containing Ψ are the derivatives with respect to ρ , ρ_s , and ρ_t . Due to (6.13) and (6.14), these remaining terms yield (6.9).

On collecting the terms proportional to Σ , we notice another cancellation. As is evident already from (6.13) and (6.14), all the terms involving cross products with respect to ρ will cancel, as they will each be multiplied by the left-hand side of (6.9) which vanishes. Thus, derivatives with respect to ρ , ρ_s , and ρ_t will not contribute to the terms proportional to Σ , so that collecting those terms will yield exactly equation (6.8). \square

There is another approach to performing the change of variables that highlights the decoupling. The key point is that we recognize two pieces of information we know about the variations $\delta \rho$, $\delta \rho_s$, and $\delta \rho_t$. First we consider the expression for $\delta \rho$ in terms of the free variations Ψ and Σ . The relation is given by

$$\delta \rho = \Psi - \Sigma \times \rho.$$

This relation can be interpreted as saying that we can select any two of the variations Σ , Ψ , and $\delta \rho$ as a free variation and the third variation is then determined. We find in practice that there are quantities such as $\delta \Omega$ that only depend on Σ . Therefore any selection of free variations must include Σ . This leaves us with a

choice of Ψ or $\delta\rho$ as the choice for the second free variation. It is interesting to consider the choice of $\delta\rho$. Indeed, since we have the relations

$$\rho_s = \partial_s \rho, \quad \rho_t = \partial_t \rho,$$

we can express the variations $\delta\rho_s$ and ρ_t in terms of our free variation $\delta\rho$:

$$\delta\rho_s = \delta\partial_s \rho = \partial_s \delta\rho,$$

similarly, $\delta\rho_t = \partial_t \delta\rho$. Since $\delta\Omega$ and $\delta\omega$ only depend on Σ we have a complete description of the variations in terms of Σ and $\delta\rho$ which are given by

$$\delta\omega = \dot{\Sigma} + \omega \times \Sigma, \quad \delta\omega = \Sigma' + \Omega \times \Sigma, \quad (6.15)$$

$$\delta\rho_s = \partial_s \delta\rho, \quad \delta\rho_t = \partial_t \delta\rho, \quad (6.16)$$

which is obviously augmented by the trivial relation $\delta\rho = \delta\rho$. An alternative proof of Theorem 6.2 can now be given.

Proof. Using the variations (6.15), (6.16) we obtain, for example, the following calculation in the variational principle,

$$\left\langle \frac{\delta\bar{L}}{\delta\rho_s}, \delta\rho_s \right\rangle = \left\langle \frac{\delta\bar{L}}{\delta\rho_s}, \partial_s \delta\rho \right\rangle = - \left\langle \partial_s \frac{\delta\bar{L}}{\delta\rho_s}, \delta\rho \right\rangle.$$

The terms arising from $\delta\Omega$ and $\delta\omega$ are identical to those before and only depend on Σ . Therefore, we obtain the following equation from stationarity under the Σ variation,

$$(\partial_t + \omega \times) \frac{\delta\bar{L}}{\delta\omega} + (\partial_s + \Omega \times) \frac{\delta\bar{L}}{\delta\Omega} = 0.$$

The second equation comes from terms proportional to $\delta\rho$ which is

$$\partial_t \frac{\delta\bar{L}}{\delta\rho_t} + \partial_s \frac{\delta\bar{L}}{\delta\rho_s} - \frac{\delta\bar{L}}{\delta\rho} = 0.$$

These are the required equations in Theorem 6.2. \square

Remark 6.2. Notice that this alternative proof does not require any cancellation of terms after the equations are derived. Thus, the variations do all the work for us. This opens up an interesting question. In some sense, the choice of $\delta\rho$ as a free variation is optimal since no extra terms appear in the resulting equations of motion. We might also refer to the heavy top at this point and ask whether a similar change of variables might simplify the heavy top equations. The answer, alas, is negative but is nevertheless instructive. The crucial property that we used was to regard $\delta\rho$ as a free variation. Now, suppose we have an advected quantity $\mathbf{a} = \Lambda^{-1}\mathbf{a}_0$. This case appears in the heavy top as well as often occurring in fluid dynamics. Could we consider $\delta\mathbf{a}$ as a free variation? Unfortunately the variation $\delta\mathbf{a}$ is given by

$$\delta\mathbf{a} = -\Sigma \times \mathbf{a}.$$

Therefore $\delta\mathbf{a}$ is determined by Σ and we cannot interpret $\delta\mathbf{a}$ as a free variation. We shall investigate the geometric structure required for this approach in Section 7.

Theorem 6.3. *The variations in $\delta \rho_s$ and $\delta \rho_t$ yield dynamical equations for the nonlocal lagrangian \bar{l}_{np} in the following form*

$$\begin{aligned} (\partial_s + \mathbf{\Omega} \times) \frac{\delta \bar{l}_{np}}{\delta \mathbf{\Omega}} &= \int \left(\mathbf{Z}(s, s') + \frac{\delta U}{\delta \kappa}(s, s') \times \kappa(s, s') \right) ds' \\ \frac{\delta \bar{l}_{np}}{\delta \rho} - \partial_s \frac{\delta \bar{l}_{np}}{\delta \rho_s} &= \int \left(\xi(s, s') \frac{\delta U}{\delta \kappa}(s', s) - \frac{\delta U}{\delta \kappa}(s, s') \right) ds', \end{aligned}$$

where, as in equation (3.24), we have

$$\mathbf{Z}(s, s') = \xi^{-1}(s, s') \left(\frac{\delta \bar{U}}{\delta \xi} \right)^T(s, s') - \frac{\delta \bar{U}}{\delta \xi}(s, s') \xi(s, s').$$

Proof. The free variations are $\widehat{\Sigma}(s) = \Lambda^{-1}(s) \delta \Lambda(s)$ and $\delta \rho(s) = \delta(\Lambda^{-1}(s) \mathbf{r}(s))$, which reads, in vector notation, $\delta \rho(s) = \Psi(s) - \Sigma(s) \times \rho(s)$. Since $\Psi(s) = \Lambda^{-1}(s) \delta \mathbf{r}(s)$ is a free variation, so is $\delta \rho(s)$. Now we directly apply the variational principle to the nonlocal Lagrangian

$$\bar{l}_{np}(\rho, \rho_s, \rho_t, \mathbf{\Omega}, \omega) = \iint U(\rho, \rho_s, \rho_t, \mathbf{\Omega}, \omega) ds ds'.$$

Recall from Section 3.1 that we have the following definitions

$$\xi(s, s') = \Lambda^{-1}(s) \Lambda(s'), \quad \kappa(s, s') = \Lambda^{-1}(s) (\mathbf{r}(s') - \mathbf{r}(s)).$$

These definitions yield the following variations

$$\left. \begin{aligned} \xi^{-1}(s, s') \delta \xi(s, s') &= \widehat{\Sigma}(s') - \text{Ad}_{\xi^{-1}(s, s')} \widehat{\Sigma}(s) \\ \delta \kappa(s, s') &= (\Sigma(s') - \xi^{-1}(s, s') \Sigma(s)) \times \rho(s') \\ &\quad + \xi(s, s') \delta \rho(s') - \delta \rho(s). \end{aligned} \right\} \quad (6.17)$$

The terms in the variational principle proportional to $\Sigma(s)$ yield

$$\begin{aligned} (\partial_s + \mathbf{\Omega} \times) \frac{\delta \bar{l}_{np}}{\delta \mathbf{\Omega}} &= \int \left(\xi(s, s') \frac{\delta U}{\delta \xi}(s', s) - \frac{\delta U}{\delta \xi}(s, s') \xi^{-1}(s, s') \right. \\ &\quad \left. + \frac{\delta U}{\delta \kappa}(s, s') \times \kappa(s, s') \right) ds', \\ &= \int \left(\mathbf{Z}(s, s') + \frac{\delta U}{\delta \kappa}(s, s') \times \kappa(s, s') \right) ds'. \end{aligned}$$

The terms in the variational principle proportional to $\delta \rho(s)$ yield

$$\partial_s \frac{\delta \bar{l}_{np}}{\delta \rho_s} - \frac{\delta \bar{l}_{np}}{\delta \rho} = \int \left(\xi(s, s') \frac{\delta U}{\delta \kappa}(s', s) - \frac{\delta U}{\delta \kappa}(s, s') \right) ds'$$

which proves the theorem. \square

Thus the combined local and nonlocal equations of motion in ρ_s and ρ_t variables are

$$\begin{aligned} & (\partial_t + \omega \times) \frac{\delta \bar{l}_{\text{loc}}}{\delta \omega} + (\partial_s + \Omega \times) \frac{\delta (\bar{l}_{\text{loc}} + \bar{l}_{\text{np}})}{\delta \Omega} \\ &= \int \left(\mathbf{Z}(s, s') + \frac{\delta U}{\delta \kappa}(s, s') \times \kappa(s, s') \right) ds', \end{aligned} \quad (6.18)$$

$$\begin{aligned} & \partial_t \frac{\delta \bar{l}_{\text{loc}}}{\delta \rho_t} + \partial_s \frac{\delta (\bar{l}_{\text{loc}} + \bar{l}_{\text{np}})}{\delta \rho_s} - \frac{\delta (\bar{l}_{\text{loc}} + \bar{l}_{\text{np}})}{\delta \rho} \\ &= \int \left(\xi(s, s') \frac{\delta U}{\delta \kappa}(s', s) - \frac{\delta U}{\delta \kappa}(s, s') \right) ds'. \end{aligned} \quad (6.19)$$

Here we see that the equations (6.18) are a form of the Euler–Poincaré equations and equations (6.19) are a form of the Euler–Lagrange equations, both modified with nonlocal terms.

Remark 6.3. This change of variables is not available in the classical Kirchhoff approach because the variable ρ is absent in this classical formulation.

6.4. The general case

We now generalize the previous results to the general situation described in Section 2.4. Recall that in this case we have $(\Lambda, r) \in \mathcal{F}(\mathcal{D}, S)$, $(\Omega, \Gamma) \in \Omega^1(\mathcal{D}, \mathfrak{s})$, and $\rho \in \mathcal{F}(\mathcal{D}, E)$, where $S = \mathcal{O} \ltimes E$ is the semidirect product of a Lie group \mathcal{O} with a vector space E .

Consider the variable ρ . Recall from (5.25) that we have the kinematic equation

$$\dot{\rho} = \gamma - \omega \rho.$$

Assuming that the initial value of ρ is zero, we have

$$\mathbf{d}\rho = \mathbf{d}(\Lambda^{-1}r) = -\Lambda^{-1}\mathbf{d}\Lambda\Lambda^{-1}r + \Lambda^{-1}\mathbf{d}r = \Gamma - \Omega\rho.$$

This motivates us to define the new variables $\rho_s \in \Omega^1(\mathcal{D}, E)$ and $\rho_t \in \mathcal{F}(\mathcal{D}, E)$ which will play the role of space and time derivatives of ρ . They are naturally defined by

$$\rho_s = \Gamma - \Omega\rho, \quad \text{and} \quad \rho_t = \gamma - \omega\rho. \quad (6.20)$$

This change of variables defines a diffeomorphism from the variables $(\omega, \gamma, \Omega, \Gamma, \rho)$ to the variables $(\omega, \Omega, \rho_s, \rho_t, \rho)$, and generalizes (6.6). In terms of the new variables, the local Lagrangian is denoted by \bar{l} and we have

$$\int_{\mathcal{D}} \bar{l}(\rho, \rho_s, \rho_t, \omega, \Omega) ds = l(\omega, \gamma, \Omega, \Gamma, \rho).$$

For simplicity, we only treat the case of a local Lagrangian. There are two equivalent points of view to obtain the equations of motion in terms of \bar{l} .

The first one is to use a variational principle, as done before in the particular case of the charged strand. Using the constrained variations of ω , γ , Ω , Γ , ρ given by the affine Euler–Poincaré principle, we obtain the constrained variations

$$\begin{aligned}\delta\omega &= \dot{\Sigma} + [\omega, \Sigma], \quad \delta\Omega = \mathbf{d}\Sigma + [\Omega, \Sigma], \\ \delta\rho_t &= \dot{\Phi} - \dot{\Sigma}\rho - \Sigma\rho_t, \quad \delta\rho_s = \mathbf{d}\Phi - \mathbf{d}\Sigma\rho - \Sigma\rho_s,\end{aligned}$$

and

$$\delta\rho = \Phi - \Sigma\rho.$$

The second point of view is to compute the functional derivatives of l in terms of those of \bar{l} . We find

$$\begin{aligned}\frac{\delta l}{\delta\omega} &= \frac{\delta\bar{l}}{\delta\omega} - \rho \diamond \frac{\delta\bar{l}}{\delta\rho_t}, \quad \frac{\delta l}{\delta\Omega} = \frac{\delta\bar{l}}{\delta\Omega} - \rho \diamond \frac{\delta\bar{l}}{\delta\rho_s}, \\ \frac{\delta l}{\delta\gamma} &= \frac{\delta\bar{l}}{\delta\rho_t}, \quad \frac{\delta l}{\delta\Gamma} = \frac{\delta\bar{l}}{\delta\rho_s},\end{aligned}$$

and

$$\frac{\delta l}{\delta\rho} = \frac{\delta\bar{l}}{\delta\rho} + \Omega_i \frac{\delta\bar{l}}{\delta\rho_{s\,i}} + \omega \frac{\delta\bar{l}}{\delta\rho_t}.$$

These two ways lead to the same equations

$$\begin{cases} \left(\frac{d}{dt} - \text{ad}_\omega^* \right) \frac{\delta\bar{l}}{\delta\omega} + \text{div}^\Omega \frac{\delta\bar{l}}{\delta\Omega} = 0, \\ \frac{d}{dt} \frac{\delta\bar{l}}{\delta\rho_t} + \text{div} \frac{\delta\bar{l}}{\delta\rho_s} - \frac{\delta\bar{l}}{\delta\rho} = 0, \end{cases} \quad (6.21)$$

where $\text{div}^\Omega : \mathfrak{X}(\mathcal{D}, \mathfrak{o}^*) \rightarrow \mathcal{F}(\mathcal{D}, \mathfrak{o}^*)$ is defined by

$$\text{div}^\Omega w := \text{div} w - \text{ad}_{\Omega_i}^* w^i \in \mathcal{F}(\mathcal{D}, \mathfrak{o}^*).$$

These equations coincide with (6.8) and (6.9) in the particular case $\mathcal{D} = [0, L]$ and $S = SE(3)$. The other equations for the advected variables are computed to be

$$\begin{cases} \dot{\rho}_s + \omega\rho_s = \mathbf{d}\rho_t + \omega\mathbf{d}\rho, \\ \dot{\Omega} + \text{ad}_\omega \Omega = \mathbf{d}\omega, \\ \dot{\rho} = \rho_t. \end{cases}$$

We also know that $\mathbf{d}\rho = \rho_s$. Therefore, using the third equation, we obtain that the first equation is verified. Thus the last system can be replaced by

$$\begin{cases} \mathbf{d}\rho = \rho_s, \\ \dot{\Omega} + \text{ad}_\omega \Omega = \mathbf{d}\omega, \\ \dot{\rho} = \rho_t. \end{cases}$$

7. The bundle covariant Lagrange–Poincaré approach

In this Section we explain how the decoupled equations discussed above are covariant Lagrange–Poincaré field equations. The coordinate change is interpreted as a transformation from the affine and modified Euler–Poincaré perspectives to the covariant Lagrange–Poincaré perspective.

The field theoretical considerations in this Section work only for local Lagrangians. Since the classical infinite dimensional approach of Sections 3 and 5 applies also to Lagrangians having a nonlocal part, it is clear that an extension of the theory presented in this section to Lagrangians with nonlocal parts exists in the field theoretic framework. Although we derived these nonlocal field equations in Section 6 by a direct coordinate transformation, at present the geometric formulation of a general theory of reduction by symmetry for nonlocal Lagrangians remains unclear. Therefore, we defer to future work a development of the Lagrange–Poincaré theory of Lagrangians depending on non-local variables.

The classical Lagrange–Poincaré equations are derived on a principal fiber bundle by using a principal connection to split the configuration space into *horizontal* and *vertical* parts. Two equations result from this splitting, one horizontal and one vertical. In the case of strands we also have to take the continuous dependence of the variables on s into account. This leads us to consider a *covariant* version of the equations.

In Section 7.1, we give various geometric structures that are required to describe the covariant Lagrange–Poincaré equations. These geometric structures are introduced very effectively in the literature and the reviews of the various geometric objects are particularly based on [6,9,27].

Section 7.2 uses these geometric structures to describe the reduction of the variational principle. The reduction is geometric in nature, insofar as it arises from a precise geometric description of the reduced covariant state space. This second part closes by describing an extra integrability condition that is required to reconstruct solutions to the unreduced equations from solutions to the reduced equations. This integrability condition has a geometric interpretation in terms of the *principal curvature* of the connection used to effect the reduction.

Finally, Sections 7.3 and 7.4, respectively, describe the Kelvin–Noether Theorem, which gives qualitative understanding of the motion of the strand, and generalizations of the strand that are supported by the theory.

7.1. Covariant state space

This section describes the geometric tools used to derive the Lagrange–Poincaré field equations. *Jet bundles and holonomic jets*, which encapsulate all physical quantities needed for a description of the strand in a single geometric object, are summarized first. A *principal bundle structure* describes the way the symmetry group interacts with the covariant state space. Finally, a *principal connection* is introduced, which allows one to split the covariant state space into two complementary parts. These tools are then brought together in Section 7.2 in order to implement the reduction.

7.1.1. Jet bundles and holonomic jets We begin by introducing *jet bundles*. Essentially, a jet bundle is a generalization of a tangent bundle to the field theoretic setting. While equations on tangent bundles geometrize the analytic notion of an ODE, equations on a jet bundle provide an analogous geometrization of a PDE.

Just as a tangent bundle plays the role of the *state space* in classical Lagrangian mechanics, a jet bundle plays the role of the *covariant state space* in covariant Lagrangian mechanics. This formulation arises from the observation that it is possible to incorporate all the dynamical information of the molecular strand into a single geometric object, namely a *holonomic jet*.

The covariant perspective is reached by noting that equations (3.38) have an exchange symmetry in their s and t dependencies. Therefore, guided by the equations derived so far, we may treat s and t on an equal basis by introducing a *space-time*, $X := [0, L] \times \mathbb{R}$. The dynamical quantities are then regarded as special vector bundle maps $\lambda : TX \rightarrow TSE(3)$. Such objects may be studied by considering the trivial fiber bundle

$$\pi_{XP} : P := X \times SE(3) \rightarrow X, \quad \pi_{XP}(x, \Lambda, \mathbf{r}) := x.$$

The analogue of the state space TQ in field theory is the *first jet bundle*, J^1P , of P .

Definition 7.1. Given a locally trivial fiber bundle $\pi_{XP} : P \rightarrow X$, the *first jet bundle* $\pi : J^1P \rightarrow P$ of P is the affine bundle over P whose fiber at $p \in P$ is

$$(J^1P)_p = \{ \lambda \in L(T_xX, T_pP) \mid T\pi_{XP} \circ \lambda = \text{id}_{T_xX} \},$$

where $L(T_xX, T_pP)$ denotes linear maps $T_xX \rightarrow T_pP$ and $x = \pi_{XP}(p)$.

The space J^1P is called the *covariant state space*.

Remark 7.1. The first jet bundle, J^1P , is a natural state space since it is the analogue of the tangent bundle in the case where sections, instead of curves, are considered. In classical Lagrangian dynamics we consider tangent vectors, (q, \dot{q}) in the state space TQ . However, we could also consider maps of the form $Tq : T\mathbb{R} \rightarrow TQ$ where \mathbb{R} is time. Since Tq is a linear map on each fiber of $T\mathbb{R}$, we consider a basis of the image of each fiber given by $T_tq(1) =: (q, \dot{q})$. Note that $T_tq(a) = (q, a\dot{q})$ for all $a \in \mathbb{R}$, which is just rescaling of time viewed in a geometric way. When one considers sections defined on spacetime instead of time dependent curves, the notion of a linear map is the idea that gives the most elegant generalization, since it captures the entire dynamics, independently of which direction is chosen in spacetime.

In field theory, a natural class of sections to consider consists of *first jet extensions* of sections of P , also called *holonomic sections* of the first jet bundle.

Definition 7.2. Let $\sigma : X \rightarrow P$ be a section of P , that is, $\pi_{XP} \circ \sigma = \text{id}_X$. The *first jet extension* of σ is the map $j^1\sigma : X \rightarrow J^1P$ defined by $j^1\sigma(x) = T_x\sigma$ for all $x \in X$.

In order to see that $j^1\sigma(x) \in (J^1P)_{\sigma(x)}$, we differentiate the relation $\pi_{XP} \circ \sigma = \text{id}_X$ to find $T\pi_{XP} \circ T\sigma = \text{id}_{TX}$. This verifies that $T\sigma \in J^1P$.

Given $X = [0, L] \times \mathbb{R}$ and $P = X \times SE(3)$, any section σ reads

$$\sigma(x) = (x, \Lambda(x), \mathbf{r}(x)) \in \{x\} \times (SO(3) \otimes \mathbb{R}^3),$$

where $x := (s, t)$. Thus $\sigma(x)$ may be identified with $(\Lambda(x), \mathbf{r}(x)) \in SE(3)$. In this case we also have $(J^1P)_{\sigma(x)} \cong L(T_x X, T_{(\Lambda(x), \mathbf{r}(x))} SE(3))$. Using this identification, we can write

$$j^1\sigma(x) = T_x\sigma \cong (\Lambda(x), \mathbf{r}(x), \Lambda'(x)ds + \dot{\Lambda}(x)dt, \mathbf{r}'(x)ds + \dot{\mathbf{r}}(x)dt).$$

Note that here we have dropped the terms x and id_{TX} to ease notation.

From (2.11) we conclude that the dependent variables that occur in the unreduced Euler–Lagrange dynamics are simply components of a first jet extension of a section of π_{XP} . Therefore, holonomic jets arise as the natural geometric objects to use in order to study the molecular strand.

7.1.2. Principal bundle structures A *principal bundle* is a particularly useful type of fiber bundle that is convenient for describing the way a Lie group acts on a manifold. More technically, a *principal bundle* is a fiber bundle together with a fiber preserving action of a Lie group that is free and transitive on each fiber. These conditions mean that each fiber is diffeomorphic to the structure group and can be described by ‘translation’ using the group action. The principal bundle is completely described by specification of the fiber bundle projection and the group action, provided the required assumptions are satisfied.

Consider, as an example, the natural principal $SO(3)$ -bundle structure on $SE(3)$ given by the projection

$$\pi_{SE(3)} : SE(3) \rightarrow \mathbb{R}^3, \quad \pi_{SE(3)}(\Lambda, \mathbf{r}) = \Lambda^{-1}\mathbf{r} = \boldsymbol{\rho},$$

and the $SO(3)$ -action

$$g(\Lambda, \mathbf{r}) = (g\Lambda, g\mathbf{r}). \quad (7.1)$$

The projection $\pi_{SE(3)}$ is invariant under the group action. Also, each fiber of $\pi_{SE(3)}$ is $SO(3)$ itself, which is acted upon by group multiplication. Therefore, the action is fiber preserving and is free and transitive on each fiber. Thus, $\pi_{SE(3)}$ with this $SO(3)$ -action is a principal bundle.

Given the example above, one may construct a new principal bundle on $P := X \times SE(3)$. Observe that $SO(3)$ acts on P by left translation on the second factor,

$$g(x, \Lambda, \mathbf{r}) = (x, g\Lambda, g\mathbf{r}).$$

Therefore, one may identify a principal $SO(3)$ -bundle structure on P over $\Sigma := P/SO(3) = X \times \mathbb{R}^3$ using this action with a projection $\pi_{\Sigma, P}$ induced by $\pi_{SE(3)}$,

$$\pi_{\Sigma P} : P \rightarrow \Sigma, \quad \pi_{\Sigma P}(x, \Lambda, \mathbf{r}) = (x, \pi_{SE(3)}(\Lambda, \mathbf{r})) = (x, \boldsymbol{\rho}).$$

In addition to P being a principal bundle over Σ , we also have the trivial fiber bundle $\pi_{X\Sigma} : \Sigma \rightarrow X$ given by

$$\pi_{X\Sigma}(x, \rho) = x.$$

To summarize, we are given a principal $SO(3)$ -bundle structure $\pi_{SE(3)}$ on $SE(3)$ and a fiber bundle structure π_{XP} on P . From these we construct a new principal $SO(3)$ -bundle structure, $\pi_{\Sigma P}$, on P and a new fiber bundle structure, $\pi_{X\Sigma}$, on $\Sigma = P/SO(3)$. Note that these projections satisfy

$$\pi_{X\Sigma} \circ \pi_{\Sigma P} = \pi_{XP}. \quad (7.2)$$

The considerations above may be summarized in the following diagram:

$$\begin{array}{ccccc} SO(3) & & SO(3) & & \\ \downarrow & & \downarrow & & \\ SE(3) & \longrightarrow & P & \xrightarrow{\pi_{XP}} & X \\ \pi_{SE(3)} \downarrow & & \pi_{\Sigma P} \downarrow & & \downarrow \text{id} \\ \mathbb{R}^3 & \longrightarrow & \Sigma & \xrightarrow{\pi_{X\Sigma}} & X. \end{array}$$

The section $\sigma(x) = (x, \Lambda(x), \mathbf{r}(x))$ of $\pi_{XP} : P \rightarrow X$ induces the section $x \mapsto (x, \rho(x))$ of $\pi_{X\Sigma} : \Sigma \rightarrow X$.

The tangent lift of the $SO(3)$ action yields a free action on the jet bundle, $J^1 P$:

$$g(\Lambda, \mathbf{r}, \Lambda' ds + \dot{\Lambda} dt, \mathbf{r}' ds + \dot{\mathbf{r}} dt) = (g\Lambda, g\mathbf{r}, g\Lambda' ds + g\dot{\Lambda} dt, g\mathbf{r}' ds + g\dot{\mathbf{r}} dt).$$

This action is free because the action on $SO(3)$ is free. The assumption that the action is proper is not required since $SO(3)$ is a compact Lie group; consequently every action of $SO(3)$ is a proper action. Therefore, we find that $J^1 P$ is also a principal $SO(3)$ -bundle. In particular, $J^1 P/SO(3)$ is a manifold. To conclude, recall from Section 7.1.1 that we identified $J^1 P$ as the covariant state space. Thus, the discussion so far, reveals that the covariant state space is a principal $SO(3)$ -bundle, which gives a convenient geometric formulation of the way the symmetry group acts.

Remark 7.2. Note that at this point we could reduce by the $SO(3)$ -action on $SE(3)$ to derive Euler–Poincaré equations. The reduced variables are given as

$$\begin{aligned} \Lambda^{-1} T(\Lambda, \mathbf{r}) &= (e, \Lambda^{-1} \mathbf{r}, \Lambda^{-1} \Lambda' ds + \Lambda^{-1} \dot{\Lambda} dt, \Lambda^{-1} \mathbf{r}' ds + \Lambda^{-1} \dot{\mathbf{r}} dt) \\ &= (\rho, \boldsymbol{\Omega} ds + \boldsymbol{\omega} dt, \boldsymbol{\Gamma} ds + \boldsymbol{\gamma} dt). \end{aligned}$$

Variations of the reduced jet $\Lambda^{-1} T(\Lambda, \mathbf{r})$ read

$$\delta \Lambda^{-1} T(\Lambda, \mathbf{r}) = (\delta \rho, \delta \boldsymbol{\Omega} ds + \delta \boldsymbol{\omega} dt, \delta \boldsymbol{\Gamma} ds + \delta \boldsymbol{\gamma} dt).$$

Therefore we may form the variational principle

$$\begin{aligned}
 \delta S &= \delta \iint l \left([j^1 \sigma] \right) ds dt, \\
 &= \iint \left\langle \frac{\delta l}{\delta \Lambda^{-1} T(\Lambda, \mathbf{r})}, \delta \Lambda^{-1} T(\Lambda, \mathbf{r}) \right\rangle ds dt, \\
 &= \iint \left(\left\langle \frac{\delta l}{\delta \rho}, \delta \rho \right\rangle + \left\langle \frac{\delta l}{\delta \Omega}, \delta \Omega \right\rangle + \left\langle \frac{\delta l}{\delta \omega}, \delta \omega \right\rangle \right. \\
 &\quad \left. + \left\langle \frac{\delta l}{\delta \Gamma}, \delta \Gamma \right\rangle + \left\langle \frac{\delta l}{\delta \gamma}, \delta \gamma \right\rangle \right) ds dt = 0.
 \end{aligned}$$

This variational principle corresponds precisely to the Euler–Poincaré variational principle from Section 3.1 with nonlocal terms removed from the Lagrangian. The derivation of the equations of motion then proceeds exactly as in Section 3.1 and results in equations (3.23) and (3.26) but with the nonlocal terms removed.

7.1.3. Principal connection A *principal connection* is a geometric tool that is required to split the space tangent to a principal bundle into horizontal and vertical parts. We introduce here the notion of a principal connection and discuss the geometric structure that it induces on P . Recall that a principal connection on a principal G -bundle P is a \mathfrak{g} -valued one form on P that satisfies

$$A(\xi_P(p)) = \xi, \quad A(gv_p) = \text{Ad}_g A(p),$$

where gv_p denotes the tangent lifted action of G on TP and ξ_P is the infinitesimal generator

$$\xi_P(p) = \left. \frac{d}{dt} \right|_{t=0} \exp(t\xi)p.$$

For our particular $SO(3)$ -bundle $\pi_{\Sigma P} : P \rightarrow \Sigma$, we make the choice

$$A(x, \Lambda, \mathbf{r}, v_x, v_\Lambda, \mathbf{u}) = v_\Lambda \Lambda^{-1} \in \mathfrak{so}(3), \quad (7.3)$$

for all $v_\Lambda \in T_\Lambda SO(3)$, $\mathbf{u} \in \mathbb{R}^3$, $v_x \in T_x X$. This connection is called the *Maurer–Cartan connection*. In general, one may choose the connection arbitrarily, but the particular choice above is well suited to the problem since it is natural. Recall that any vector $v_\Lambda \in T_\Lambda SO(3)$ may be written $v_\Lambda = \Lambda \eta$ where $\eta \in \mathfrak{so}(3)$. The connection decomposes TP into the *horizontal* and *vertical subbundles* as follows:

$$\text{Ver } P = \ker(T\pi_{\Sigma P}) = \{(x, \Lambda, \mathbf{r}; 0, \Lambda \eta, (\text{Ad}_\Lambda \eta) \mathbf{r}) \mid \eta \in \mathfrak{so}(3)\}, \quad (7.4)$$

$$\text{Hor}_A P = \ker A = \{(x, \Lambda, \mathbf{r}; v_x, 0, \mathbf{u}) \mid \mathbf{u} \in T_{\mathbf{r}} \mathbb{R}^3, v_x \in T_x X\}. \quad (7.5)$$

These two subbundles are complementary in the sense that

$$TP = \text{Ver } P \oplus \text{Hor}_A P,$$

where \oplus denotes the Whitney sum of vector bundles. This geometric decomposition of TP into complementary subbundles induces a similar decomposition of the covariant state space $J^1 P$ and ultimately results in two Lagrange–Poincaré equations, one for each factor of the decomposition.

7.2. The variational principle

Having introduced the geometric setting for the Lagrange–Poincaré field approach, we formulate the reduced variational principle and derive the associated field equations of motion.

The formulation of the reduced variational principle results from geometric properties of the *reduced state space* $J^1P/SO(3)$. To understand it, we begin with the study of the simpler reduced bundle $TP/SO(3)$ in Section 7.2.1.

The investigations in Section 7.2.1 lead to the introduction of an *associated bundle*, namely the *adjoint bundle*, which describes the vertical part of the reduced bundle $TP/SO(3)$. The properties of the adjoint bundle are reviewed in Section 7.2.2.

In Section 7.2.3 we generalize these considerations to the covariant setting and provide a discussion on the geometric structure of J^1P/G .

Having gained a sufficient understanding of the geometry behind the reduced state space J^1P/G , we turn to deriving the reduced variations in Section 7.2.4 and then use these variations in the formulation of the reduced variational principle in Section 7.2.5.

In Section 7.2.6 we discuss and present the reconstruction of solutions to the unreduced equations from those of the reduced equations.

7.2.1. Splitting $TP/SO(3)$ The principal bundle structure $\pi_{\Sigma,P} : P \rightarrow \Sigma$ introduces the geometric tools to formally decompose the variational principle on $TP/SO(3)$ into two parts that correspond to two free variations. While in the present situation it is possible to write down such variations globally on $TP/SO(3)$, in general, it is not possible to do this. Such a formulation only makes sense locally in a trivialization of $\pi_{\Sigma,P}$. In addition, a particular choice of gauge must be chosen on P . For a general formulation of the variations one requires a global geometric splitting of $TP/SO(3)$ in terms of a connection form. This splitting takes the form of a bundle isomorphism $\alpha_A : TP/SO(3) \rightarrow T\Sigma \oplus \text{Ad}P$ where \oplus denotes the Whitney sum of vector bundles over Σ and $\text{Ad}P$ is a vector bundle over Σ introduced below. The components of the Whitney sum correspond to the two free variations. The bundle isomorphism is given by

$$\alpha_A \left([v_p]_{SO(3)} \right) := T\pi_{\Sigma P}(v_p) \oplus [p, A(v_p)]_{SO(3)}, \quad v_p \in T_pP,$$

where $\text{Ad}P = (P \times \mathfrak{so}(3))/SO(3)$ is the *adjoint bundle* to P . The quotient is taken relative to the left diagonal action

$$(p, \eta) \mapsto (hp, \text{Ad}_h \eta),$$

and the elements in the adjoint bundle are written as $[p, \eta]_{SO(3)}$. To check that α_A is well defined one easily verifies that

$$T\pi_{\Sigma P}(hv_p) = T\pi_{\Sigma P}(v_p),$$

and

$$[hp, A(hv_p)]_{SO(3)} = [hp, \text{Ad}_h A(v_p)]_{SO(3)} = [p, A(v_p)]_{SO(3)}.$$

To show that α_A is an isomorphism, we give its inverse

$$\alpha_A^{-1}(v_{(x,\rho)} \oplus [p, \eta]_{SO(3)}) = [\text{Hor}_p^A(v_{(x,\rho)}) + \eta_P(p)]_{SO(3)},$$

where $p = (x, \Lambda, \mathbf{r}) \in P$ is such that $\pi_{\Sigma P}(p) = (x, \rho)$ and Hor_p^A denotes the *horizontal lift* of $v_{(x,\rho)} = (x, \rho, v_x, \mathbf{u}) \in T_{(x,\rho)}\Sigma$ to $T_p P$ with respect to A . That is, $\text{Hor}_p^A : T_\rho \Sigma \rightarrow H_p P$ is the unique map such that

$$T_p \pi_{\Sigma, P} \circ \text{Hor}_p^A = \text{id}_{T_\rho \Sigma}.$$

The horizontal lift is given explicitly by

$$\text{Hor}_{(x,\Lambda,\mathbf{r})}^A(x, \rho, v_x, \mathbf{u}) = (x, \Lambda, \mathbf{r}, v_x, 0, \Lambda \mathbf{u}).$$

Remark 7.3. (The choice of connection) As we have seen in (7.3), a natural choice of connection is the Maurer–Cartan form $v_\Lambda \mapsto v_\Lambda \Lambda^{-1}$.

7.2.2. Properties of $\text{Ad}P$ Various properties of the adjoint bundle are required to derive the Lagrange–Poincaré equations; we now review these properties.

We can give $\text{Ad}P$ a Lie algebra structure on each fiber. The vector space structure is given by

$$[p, \eta]_{SO(3)} + a[p, v]_{SO(3)} = [p, \eta + av]_{SO(3)}$$

and the Lie bracket is given by

$$[[p, \eta]_{SO(3)}, [p, v]_{SO(3)}] = [p, [\eta, v]]_{SO(3)}.$$

The principal connection A induces an affine connection on the adjoint bundle $\text{Ad}P$. It is known that the covariant derivative of this affine connection is given by

$$\frac{D^A}{D\tau} [p(\tau), \eta(\tau)]_{SO(3)} = [p(\tau), \dot{\eta}(\tau) - [A(\dot{p}(\tau)), \eta(\tau)]]_{SO(3)} \quad (7.6)$$

(see, for example, [6], Lemma 2.3.4).

This formula allows us to define a covariant derivative of any smooth map $\zeta : X \rightarrow \text{Ad}P$ by using the formula for the covariant derivative of the vector bundle $\text{Ad}P \rightarrow \Sigma$ induced by the principal connection A on $\pi_{\Sigma P} : P \rightarrow \Sigma$. We define

$$\nabla_U^A \zeta(x) := \left. \frac{D^A}{D\tau} \right|_{\tau=0} (\zeta \circ c)(\tau), \quad (7.7)$$

where $c(\tau)$ is a smooth curve in X such that $c(0) = x$ and $\dot{c}(0) = U \in T_x X$. Concretely, denoting $\zeta(x) = [p(x), \eta(x)]_{SO(3)}$, formula (7.6) gives

$$\nabla_U^A \zeta(x) = [p(x), \mathbf{d}\eta(x)(U) - [A(T_x p(U)), \eta(x)]]_{SO(3)}. \quad (7.8)$$

See [10] for further details on the construction of these differential operators. Note that the map ζ is not a section of the adjoint bundle since ζ is defined on X , whereas the adjoint bundle has base Σ .

Note that the vector bundle $\text{Ad}P \rightarrow \Sigma$ is trivial in the present case. Indeed, the map

$$[(x, \Lambda, \mathbf{r}), \eta]_{SO(3)} \mapsto \left((x, \Lambda^{-1}\mathbf{r}), \text{Ad}_{\Lambda^{-1}} \eta \right) \quad (7.9)$$

is a vector bundle isomorphism from $\text{Ad}P$ to $\Sigma \times \mathfrak{so}(3)$. In this trivialization, using the connection (7.3), the formula for the covariant derivative (7.6) becomes

$$\frac{D^A}{D\tau}(x(\tau), \boldsymbol{\rho}(\tau), \xi(\tau)) = (x(\tau), \boldsymbol{\rho}(\tau), \dot{\xi}(\tau)).$$

Similarly, if $U \in T_x X$, formula (7.8) becomes

$$\nabla_U^A \zeta(x) = (x, \boldsymbol{\rho}(x), \mathbf{d}\xi(x)(U)), \quad \text{where } \zeta(x) = (x, \boldsymbol{\rho}(x), \xi(x)). \quad (7.10)$$

Had the bundles been nontrivial, the formulas for the covariant derivatives would be more involved.

7.2.3. Splitting $J^1 P/SO(3)$ Having introduced the connection that splits $TP/SO(3)$, we now wish to use it to split the reduced covariant state space $J^1 P/SO(3)$. This is easily achieved by regarding the jets as linear maps and composing with α_A . Therefore, we split $J^1 P/SO(3)$ by splitting the image of the jets in $TP/SO(3)$.

Consider a section σ of the fiber bundle $\pi_{XP} : P \rightarrow X$ and its first jet extension $j^1\sigma(x) = T_x\sigma \in (J^1 P)_{\sigma(x)}$. Composing $[T\sigma]_{SO(3)} : TX \rightarrow TP/SO(3)$ with the vector bundle isomorphism $\alpha_A : TP/SO(3) \rightarrow T\Sigma \oplus_\Sigma \text{Ad}P$ over Σ yields the following equality in the fiber over $\pi_{\Sigma P}(\sigma(x))$:

$$\begin{aligned} \alpha_A \circ [T_x\sigma]_{SO(3)} &= (T_{\sigma(x)}\pi_{\Sigma P} \circ T_x\sigma) \oplus [\sigma(x), A \circ T_x\sigma]_{SO(3)} \\ &= T_x(\pi_{\Sigma P} \circ \sigma) \oplus [\sigma(x), A \circ T_x\sigma]_{SO(3)}. \end{aligned}$$

Using $\pi_{X\Sigma} \circ \pi_{\Sigma P} = \pi_{XP}$, we have

$$\pi_{X\Sigma} \circ (\pi_{\Sigma P} \circ \sigma) = \pi_{XP} \circ \sigma = \text{id}_X.$$

This shows that $\pi_{\Sigma P} \circ \sigma$ is a section of the fiber bundle $\pi_{X\Sigma} : \Sigma \rightarrow X$. If we denote

$$\sigma_1 = \pi_{\Sigma P} \circ \sigma, \quad \bar{A}(v_p) := [p, A(v_p)]_{SO(3)}, \quad \text{and} \quad \sigma_2(x) := \bar{A} \circ T_x\sigma,$$

for all $v_p \in T_p P$, then the *reduced jet* $[j^1\sigma(x)]_{SO(3)} \in (J^1 P)/SO(3)$ may be expressed as

$$\alpha_A \circ [j^1\sigma(x)]_{SO(3)} = \alpha_A \circ [T_x\sigma]_{SO(3)} = T_x\sigma_1 \oplus \bar{A} \circ T_x\sigma = T_x\boldsymbol{\rho} \oplus \sigma_2(x),$$

since $\sigma_1 = \boldsymbol{\rho}$. Note that this element lies in the fiber

$$(J^1\Sigma)_{\sigma_1(x)} \times L(T_x X, (\text{Ad}P)_{\sigma_1(x)})$$

over $\sigma_1(x) = \pi_{X\Sigma}(\sigma(x)) \in \Sigma$. In particular, there is a fiber bundle isomorphism

$$J^1 P / SO(3) \cong J^1 \Sigma \times_{\Sigma} L(TX, \text{Ad}P)$$

over Σ . Using the equality $\sigma(x) = (x, \Lambda(x), \mathbf{r}(x))$, the explicit description of the quantities appearing in the reduced jet are:

$$\begin{aligned} T\rho &= \left(\Lambda^{-1} \mathbf{r}, d(\Lambda^{-1} \mathbf{r}) \right) = (\rho, \rho_s ds + \rho_t dt) \in \rho^*(J^1 \Sigma) \\ \sigma_2 &= \left[(\Lambda, \mathbf{r}), d\Lambda \Lambda^{-1} \right]_{SO(3)} \cong \left((\Lambda^{-1} \mathbf{r}), \Lambda^{-1} d\Lambda \right) = (\rho, \Omega ds + \omega dt), \end{aligned}$$

by the relations (2.11), where \cong denotes here the vector bundle isomorphism (7.9). Thus, the reduced jet $[j^1 \sigma(x)]_{SO(3)}$ associated to $\sigma(x) = (x, \Lambda(x), \mathbf{r}(x))$ is represented in the trivialization (7.9) by

$$\begin{aligned} T\rho \oplus \sigma_2 &= \left(\Lambda^{-1} \mathbf{r}, d(\Lambda^{-1} \mathbf{r}), \Lambda^{-1} d\Lambda \right) \\ &= (\rho, \rho_s ds + \rho_t dt, \Omega ds + \omega dt). \end{aligned} \quad (7.11)$$

Therefore we have recovered the new coordinates given in (6.5) and the reduced Lagrangian reads

$$\bar{l}(T\rho, \sigma_2) = \bar{l}(\rho, \rho_s, \rho_t, \omega, \Omega).$$

The spaces involved in the previous discussion are summarized in the following diagram:

$$\begin{array}{ccccc} \text{Ad}SE(3) & & \text{Ad}P & & \\ \downarrow & & \downarrow & & \\ TSE(3)/SO(3) & \longrightarrow & TP/SO(3) & \xrightarrow{T\pi_{XP}/SO(3)} & TX \\ T\pi_{SE(3)}/SO(3) \downarrow & & T\pi_{\Sigma P}/SO(3) \downarrow & & \downarrow \text{id} \\ T\mathbb{R}^3 & \longrightarrow & T\Sigma & \xrightarrow{T\pi_{X\Sigma}} & TX. \end{array}$$

Recall that the projections π_{XP} and $\pi_{\Sigma P}$ are $SO(3)$ -invariant, therefore they naturally induce projections $\pi_{\Sigma P}/SO(3) : P/SO(3) \rightarrow \Sigma$ and $\pi_{XP}/SO(3) : P/SO(3) \rightarrow X$, respectively. The relationship between the variables in the affine Euler–Poincaré and covariant Lagrange–Poincaré equations can be illustrated in the following diagram:

$$\begin{array}{ccccc} (\rho, \Omega ds + \omega dt)_{SO(3)} \in L(TX, \text{Ad}SE(3)) & & L(TX, \text{Ad}P) & & \\ \downarrow & & \downarrow & & \\ (\rho, \Gamma ds + \gamma dt, \Omega ds + \omega dt) \in L(TX, TSE(3)/SO(3)) & \longrightarrow & J^1 P / SO(3) & \xrightarrow{T\pi_{XP}/SO(3)} & L(TX, TX) \\ T\pi_{SE(3)}/SO(3) \downarrow & & T\pi_{\Sigma P}/SO(3) \downarrow & & \downarrow \text{id} \\ (\rho, \rho_s ds + \rho_t dt) \in L(TX, T\mathbb{R}^3) & \longrightarrow & J^1 \Sigma & \xrightarrow{T\pi_{X\Sigma}} & L(TX, TX). \end{array}$$

The affine Euler–Poincaré variables $[j^1 \sigma]_{SO(3)} \cong (\rho, \Omega ds + \omega dt, \Gamma ds + \gamma dt)$ appear in the middle horizontal sequence whereas the covariant Lagrange–Poincaré variables $(\rho, \rho_s ds + \rho_t dt, \Omega ds + \omega dt)$ appear in the top and bottom rows.

7.2.4. Reduced variations Having split the reduced state space $J^1P/SO(3)$ into horizontal and vertical parts, we now proceed to calculate the reduced variations in each of these factors.

Let $\sigma : X \rightarrow P$ be a section of the fiber bundle $\pi_{XP} : P \rightarrow X$. If $\sigma_\varepsilon : X \rightarrow P$ is a curve of sections with $\sigma_0 = \sigma$, that is, $\sigma_\varepsilon(x) = (x, \Lambda_\varepsilon(x), \mathbf{r}_\varepsilon(x))$, $\Lambda_0 = \Lambda$, and $\mathbf{r}_0 = \mathbf{r}$, define the variation

$$\delta\sigma(x) = \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} (x, \Lambda_\varepsilon(x), \mathbf{r}_\varepsilon(x)) \in T_{\sigma(x)}P.$$

Splitting $\delta\sigma(x)$ into its vertical and horizontal parts relative to the connection A in the principal $SO(3)$ -bundle $\pi_{\Sigma P} : P \rightarrow \Sigma$ (see (7.4), (7.5)) gives

$$\begin{aligned} \delta\sigma(x) &= (x, \Lambda, \mathbf{r}, 0, \delta\Lambda, \delta\mathbf{r}) \\ &= \left(x, \Lambda, \mathbf{r}, 0, \delta\Lambda, \delta\Lambda\Lambda^{-1}\mathbf{r} \right) + \left(x, \Lambda, \mathbf{r}, 0, 0, \delta\mathbf{r} - \delta\Lambda\Lambda^{-1}\mathbf{r} \right) \in T_{\sigma(x)}P. \end{aligned}$$

To compute the vertical variation of $[j^1\sigma]_{SO(3)}$, we consider curves σ_ε that perturb $\sigma_0 = \sigma : X \rightarrow P$ along the group orbits, that is,

$$\sigma_\varepsilon(x) := \exp(\varepsilon\xi(x)) \cdot \sigma(x) = (x, \exp(\varepsilon\xi(x))\Lambda(x), \exp(\varepsilon\xi(x))\mathbf{r}(x)),$$

where $\xi : X \rightarrow \mathfrak{so}(3)$. By (7.11), in the trivialization (7.9),

$$\left[j^1\sigma_\varepsilon \right]_{SO(3)} \cong \left(\Lambda^{-1}\mathbf{r}; d(\Lambda^{-1}\mathbf{r}), (\exp(\varepsilon\xi)\Lambda)^{-1} d(\exp(\varepsilon\xi)\Lambda) \right).$$

Taking the ε -derivative of the right-hand side we get the vertical variation

$$\begin{aligned} \delta^v \left[j^1\sigma \right]_{SO(3)} &:= \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \left[j^1\sigma_\varepsilon \right]_{SO(3)} \\ &\cong \left(\Lambda^{-1}\mathbf{r}; 0, (\boldsymbol{\Sigma}' + \boldsymbol{\Omega} \times \boldsymbol{\Sigma}) ds + (\widehat{\boldsymbol{\Sigma}} + \boldsymbol{\omega} \times \boldsymbol{\Sigma}) dt \right), \quad (7.12) \end{aligned}$$

where $\widehat{\boldsymbol{\Sigma}}(x) := \text{Ad}_{\Lambda(x)^{-1}} \xi(x)$.

To compute the horizontal variation of $[j^1\sigma]_{SO(3)}$, we consider curves σ_ε that perturb $\sigma_0 = \sigma : X \rightarrow P$ such that $\delta\sigma$ is horizontal. In view of (7.5), a curve giving a horizontal $\delta\sigma$ is $\sigma_\varepsilon(x) := (x, \Lambda(x), \mathbf{r}_\varepsilon(x))$. Therefore, for such a curve σ_ε we get

$$\left[j^1\sigma_\varepsilon \right]_{SO(3)} \cong \left(\Lambda^{-1}\mathbf{r}_\varepsilon; d(\Lambda^{-1}\mathbf{r}_\varepsilon), \Lambda^{-1}d\Lambda \right)$$

and hence the horizontal variation is

$$\delta^h \left[j^1\sigma \right]_{SO(3)} := \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \left[j^1\sigma_\varepsilon \right]_{SO(3)} \cong \left(\Lambda^{-1}\mathbf{r}; d(\Lambda^{-1}\delta\mathbf{r}), 0 \right) \quad (7.13)$$

$$= (\boldsymbol{\rho}; d(\delta\boldsymbol{\rho}), 0) = (\boldsymbol{\rho}; \partial_t(\delta\boldsymbol{\rho}) dt + \partial_s(\delta\boldsymbol{\rho}) ds, 0). \quad (7.14)$$

Remark 7.4. The free variations $\delta\boldsymbol{\rho}$ and $\boldsymbol{\Sigma}$ are now recognized as being horizontal and vertical variations. This is the reason for the decoupled form of the resulting equations. If we had defects in the strand and therefore our connection had non-zero curvature, then the equations would not decouple completely. See Section 7.2.6 for more details on curvature.

7.2.5. The variational principle Having derived the reduced horizontal and vertical variations, we may now derive the horizontal and vertical Lagrange–Poincaré equations. For the vertical variations, using (7.12) and

$$\frac{\delta \bar{l}}{\delta \sigma_2} = \left[\rho, \frac{\delta \bar{l}}{\delta \omega} \partial_t + \frac{\delta \bar{l}}{\delta \Omega} \partial_s \right]_{SO(3)},$$

we obtain

$$\begin{aligned} \delta^v S &= \frac{d}{d\varepsilon} \Big|_{\varepsilon=0} \int_X \bar{l} \left([j^1 \sigma_\varepsilon]_{SO(3)} \right) dx \\ &= \int_X \left\langle \left[\rho, \frac{\delta \bar{l}}{\delta \omega} \partial_t + \frac{\delta \bar{l}}{\delta \Omega} \partial_s \right]_{SO(3)}, [\rho, (\dot{\Sigma} + \omega \times \Sigma) dt + (\Sigma' + \Omega \times \Sigma) ds]_{SO(3)} \right\rangle dx \\ &= - \int_X \left\langle \left[\rho, (\partial_t + \omega \times) \frac{\delta \bar{l}}{\delta \omega} + (\partial_s + \Omega \times) \frac{\delta \bar{l}}{\delta \Omega} \right]_{SO(3)}, [\rho, \Sigma]_{SO(3)} \right\rangle dx = 0. \end{aligned}$$

Therefore, the *vertical covariant Lagrange–Poincaré equation* is

$$(\partial_t + \omega \times) \frac{\delta \bar{l}}{\delta \omega} + (\partial_s + \Omega \times) \frac{\delta \bar{l}}{\delta \Omega} = 0.$$

Similarly, we derive the variational principle for horizontal variations and obtain, using (7.13),

$$\begin{aligned} \delta^h S &= \frac{d}{d\varepsilon} \Big|_{\varepsilon=0} \int_X \bar{l} \left([j^1 \sigma_\varepsilon]_{SO(3)} \right) dx = \int_X \left\langle \frac{\delta \bar{l}}{\delta T \rho}, \delta T \rho \right\rangle dx \\ &= \int_X \left\langle \frac{\delta \bar{l}}{\delta \rho}, \delta \rho \right\rangle + \left\langle \frac{\delta \bar{l}}{\delta \rho_t} \partial_t + \frac{\delta \bar{l}}{\delta \rho_s} \partial_s, \delta \rho_t dt + \delta \rho_s ds \right\rangle dx \\ &= \int_X \left\langle \frac{\delta \bar{l}}{\delta \rho} - \partial_t \frac{\delta \bar{l}}{\delta \rho_t} - \partial_s \frac{\delta \bar{l}}{\delta \rho_s}, \delta \rho \right\rangle dx = 0. \end{aligned}$$

Therefore, the *horizontal covariant Lagrange–Poincaré equation* is

$$\frac{\delta \bar{l}}{\delta \rho} - \partial_t \frac{\delta \bar{l}}{\delta \rho_t} - \partial_s \frac{\delta \bar{l}}{\delta \rho_s} = 0.$$

Upon putting these together, we find that the covariant Lagrange–Poincaré equations are

$$(\partial_t + \omega \times) \frac{\delta \bar{l}}{\delta \omega} + (\partial_s + \Omega \times) \frac{\delta \bar{l}}{\delta \Omega} = 0, \quad (7.15)$$

$$\frac{\delta \bar{l}}{\delta \rho} - \partial_t \frac{\delta \bar{l}}{\delta \rho_t} - \partial_s \frac{\delta \bar{l}}{\delta \rho_s} = 0. \quad (7.16)$$

Now we see that equations (7.15) and (7.16) are, respectively, identical to equations (6.8) and (6.9) which arise as the local part of equations (3.23) and (3.26), respectively, in terms of the change of variables described in equation (6.6). Thus, the natural interpretation of the change of variables in Section 6 is that they transform from the affine Euler–Poincaré formulation described in Section 5 into a form

of Lagrange–Poincaré field theory that has been modified by nonlocal terms. At present, there is no general theory for nonlocal Lagrange–Poincaré field reduction. The development of such a theory would be very interesting. However, it is beyond the scope of the present paper. For a full discussion and development of the general theory of local Lagrange–Poincaré field reduction with applications, see [10].

7.2.6. Reconstruction and curvature relations In [10] and [27] it is noted that for Lagrange–Poincaré field reduction an extra integrability condition is required to reconstruct the solution to the original Euler–Lagrange equations from the Lagrange–Poincaré field equations. This extra condition involves the curvature form of the principal connection A used for Lagrange–Poincaré field reduction.

Physically, the reconstruction condition may be interpreted as containing information about topological defects in the orientation field (the $SO(3)$ field). This interpretation arises in nematic liquid crystals (see, for example, [18]), where integrals of curvature forms provide information on defects in the liquid crystal. Remarkably, these integrals are related to characteristic classes of the principal bundles involved and, consequently, contain topological information about the principal bundle expressed in terms of the curvature of the connection.

Geometrically, the vanishing of the curvature of the connection one-form A is equivalent to integrability of the horizontal subbundle $\text{Hor}_A P$ by the Frobenius Theorem (see, for example, [6]). Therefore, when the curvature vanishes, that is, one has a zero-curvature relation, one may conclude that there are horizontal submanifolds whose tangent space at a point $p \in P$ is the horizontal space $(\text{Hor}_A P)_p$. This property of the curvature of a principal connection describes in a geometrical way the necessary integrability condition for reconstruction; see [10] and [27].

In the current setting, our principal connection A (the Maurer–Cartan connection) has zero-curvature since the following equation is satisfied,

$$B(u_p, v_p) = \mathbf{d}A(u_p, v_p) - [A(u_p)A(v_p)] = 0, \quad (7.17)$$

where B denotes the curvature 2-form of A and $[\cdot, \cdot]$ denotes the bracket on the Lie algebra. That is, the orientations mounted on the filament do not possess any topological defects and the horizontal bundle is integrable. Since $A(\Lambda, \mathbf{r}, v_\Lambda \Lambda^{-1}, v_{\mathbf{r}}) = v_\Lambda \Lambda^{-1}$ one sees that equation (7.17) reads

$$B((\Lambda, \mathbf{r}, \dot{\Lambda}, \dot{\mathbf{r}}), (\Lambda, \mathbf{r}, \Lambda', \mathbf{r}')) = \text{Ad}_\Lambda (\partial_t \boldsymbol{\Omega} - \partial_s \boldsymbol{\omega} - \boldsymbol{\omega} \times \boldsymbol{\Omega})^\wedge = 0.$$

Thus the zero-curvature integrability condition required for reconstruction is

$$\partial_t \boldsymbol{\Omega} - \partial_s \boldsymbol{\omega} - \boldsymbol{\omega} \times \boldsymbol{\Omega} = 0. \quad (7.18)$$

This equation was previously derived directly from equality of cross-derivatives and stated in (2.19), where it was used to close the dynamics that arose from the variational principle. Now we recognize that these equations assert an integrability condition that allows reconstruction of the full dynamics from the reduced system.

7.3. The Kelvin–Noether Theorem

The Kelvin–Noether Theorem gives qualitative information about the behavior of solutions to the Euler–Poincaré equations in continuum mechanics. There is an analogue of the Kelvin–Noether Theorem in the covariant picture, which we describe below.

Denoting by div_x the divergence relative to the variable $x = (s, t) \in [0, L] \times \mathbb{R}$, we have

$$\begin{aligned} \text{div}_x \left(\text{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}}{\delta \bar{\Omega}} \partial_s + \text{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}}{\delta \bar{\omega}} \partial_t \right) \\ = \partial_s \left(\text{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}}{\delta \bar{\Omega}} \right) + \partial_t \left(\text{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}}{\delta \bar{\omega}} \right) \\ = \text{Ad}_{\Lambda^{-1}}^* \left(\partial_s \frac{\delta \bar{l}}{\delta \bar{\Omega}} + \bar{\Omega} \times \frac{\delta \bar{l}}{\delta \bar{\Omega}} + \partial_t \frac{\delta \bar{l}}{\delta \bar{\omega}} + \bar{\omega} \times \frac{\delta \bar{l}}{\delta \bar{\omega}} \right) = 0 \end{aligned}$$

by (7.15). Therefore, we gain the following qualitative information: when the strand is a closed loop, the change of variables (1.15) implies that the circulation of its spatial angular momentum integrated around the moving loop in the convective frame is conserved. That is,

$$\frac{d}{dt} \oint \left(\text{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}}{\delta \bar{\Omega}} \right) ds = - \oint \partial_s \left(\text{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}}{\delta \bar{\Omega}} \right) ds = 0. \quad (7.19)$$

By using the divergence theorem, this circulation theorem may be re-expressed covariantly as

$$\begin{aligned} 0 &= \int_S \text{div}_x \left(\text{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}}{\delta \bar{\Omega}} \partial_s + \text{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}}{\delta \bar{\omega}} \partial_t \right) ds dt \\ &= \int_{\partial S} \left(\text{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}}{\delta \bar{\Omega}} \partial_s + \text{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}}{\delta \bar{\omega}} \partial_t \right) \cdot \mathbf{n} d\ell \\ &= \int_{\partial S} \text{Ad}_{\Lambda^{-1}}^* \left(\frac{\delta \bar{l}}{\delta \bar{\Omega}} dt - \frac{\delta \bar{l}}{\delta \bar{\omega}} ds \right), \end{aligned}$$

where \mathbf{n} is the outward pointing unit normal to the boundary ∂S and we have used the identity

$$\left(\frac{\delta \bar{l}}{\delta \bar{\Omega}} \partial_s + \frac{\delta \bar{l}}{\delta \bar{\omega}} \partial_t \right) \cdot \mathbf{n} d\ell = \frac{\delta \bar{l}}{\delta \bar{\Omega}} dt - \frac{\delta \bar{l}}{\delta \bar{\omega}} ds. \quad (7.20)$$

Thus, the covariant expression of the circulation theorem (7.19) becomes

$$\int_{\partial S} \text{Ad}_{\Lambda^{-1}}^* \left(\frac{\delta \bar{l}}{\delta \bar{\Omega}} dt - \frac{\delta \bar{l}}{\delta \bar{\omega}} ds \right) = 0, \quad (7.21)$$

which may be interpreted as a zero-flux theorem in ‘spacetime’ X .

7.4. Generalizations of the molecular strand

The covariant Lagrange–Poincaré approach generalizes to the case of an n -dimensional strand with an arbitrary Lie group structure \mathcal{O} , in the setting of Section 2.4, as follows. Consider the $(n + 1)$ -dimensional spacetime $X := \mathcal{D} \times \mathbb{R}$ and the trivial fiber bundle

$$\pi_{XP} : P := X \times S \rightarrow X,$$

where $S = \mathcal{O} \otimes E$ and E is a representation space of \mathcal{O} . A section σ of P reads

$$\sigma(x) = (x, \Lambda(x), r(x)), \quad x = (s, t) \in X,$$

and its first jet extension is

$$j^1\sigma(x) = (T_x\Lambda, T_xr) = (\mathbf{d}\Lambda(x) + \dot{\Lambda}(x)dt, \mathbf{d}r(x) + \dot{r}(x)dt),$$

where \mathbf{d} is the partial derivative with respect to space (that is, the derivative on \mathcal{D}), and the dot is the partial derivative with respect to time.

There is a natural \mathcal{O} -principal bundle structure on S given by

$$\pi_{ES} : S \rightarrow E, \quad \pi_{ES}(\Lambda, r) = \Lambda^{-1}r = \rho.$$

This principal bundle structure on the fiber S induces a principal \mathcal{O} -bundle structure on P given by

$$\pi_{\Sigma P} : P \rightarrow X \times E, \quad \pi_{\Sigma P}(x, \Lambda, r) = (x, \Lambda^{-1}r).$$

There is a natural connection A on $\pi_{\Sigma P} : P \rightarrow \Sigma := X \times E$ given by

$$A(v_x, v_\Lambda, (r, u)) = v_\Lambda \Lambda^{-1},$$

which allows us to identify the reduced jet bundle J^1P/\mathcal{O} with the fiber bundle $J^1\Sigma \times_\Sigma L(TX, \text{Ad}P)$. Using the same notations as before, we have

$$\alpha_A \circ [j^1\sigma(x)]_{\mathcal{O}} = T_x\rho \oplus \bar{A} \circ T_x\sigma \cong (x, \rho(x), \mathbf{d}\rho(x) + \dot{\rho}(x), \Omega(x) + \omega(x)dt),$$

by (2.20). The vertical and horizontal variations being given by

$$\begin{aligned} \delta^v \left[j^1\sigma(x) \right]_{\mathcal{O}} &= (x, \Lambda^{-1}r(x); 0, \mathbf{d}\Sigma + [\Omega, \Sigma] + (\dot{\Sigma} + [\omega, \Sigma])dt), \\ \delta^h \left[j^1\sigma(x) \right]_{\mathcal{O}} &= (x, \rho(x); \mathbf{d}(\delta\rho) + (\delta\rho)_t dt, 0), \end{aligned}$$

we find that the vertical and horizontal Lagrange–Poincaré equations are

$$\begin{aligned} (\partial_t - \text{ad}_\omega^*) \frac{\delta \bar{l}}{\delta \omega} + (\text{div} - \text{ad}_\Omega^*) \frac{\delta \bar{l}}{\delta \Omega} &= 0, \\ \frac{\delta \bar{l}}{\delta \rho} - \partial_t \frac{\delta \bar{l}}{\delta \rho_t} - \text{div} \frac{\delta \bar{l}}{\delta \rho_s} &= 0. \end{aligned}$$

Of course, as expected, these equations coincide with equations (6.21) obtained from the affine Euler–Poincaré equations (5.24) by the change of variables (6.20).

In Appendix D we derive the same equations by applying the covariant Lagrange–Poincaré approach in [27], using the group structure of S . Note that the approach we have used here does not use the group structure of S and so one may expect it to apply to more general situations such as the molecular strand on the sphere. We have

$$\begin{aligned} \operatorname{div}_x \left(\operatorname{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}}{\delta \Omega} + \operatorname{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}}{\delta \omega} \partial_t \right) &= \operatorname{div} \left(\operatorname{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}}{\delta \Omega} \right) + \partial_t \left(\operatorname{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}}{\delta \omega} \right) \\ &= \operatorname{Ad}_{\Lambda^{-1}}^* \left(\operatorname{div} \frac{\delta \bar{l}}{\delta \Omega} - \operatorname{ad}_{\Omega}^* \frac{\delta \bar{l}}{\delta \Omega} + \partial_t \frac{\delta \bar{l}}{\delta \omega} - \operatorname{ad}_{\omega}^* \frac{\delta \bar{l}}{\delta \omega} \right) = 0. \end{aligned}$$

Using the divergence theorem, we find the generalization of the zero flux theorem (7.21), now applied to an n -dimensional strand (or surface),

$$\int_{\partial V} \left(\operatorname{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}}{\delta \Omega} + \operatorname{Ad}_{\Lambda^{-1}}^* \frac{\delta \bar{l}}{\delta \omega} \partial_t \right) \cdot \mathbf{n} \, d\sigma = 0,$$

where \mathbf{n} is the outward pointing unit normal to the boundary ∂V of a given domain $V \subset \mathcal{D} \times \mathbb{R}$ and $d\sigma$ is the induced boundary volume element of ∂V .

Remark 7.5. The method used here for the particular case of $P = X \times SE(3)$ and $\Sigma = P/SO(3) = X \times \mathbb{R}^3$ can be adapted to apply more generally to a fiber bundle $\pi_{XP} : P \rightarrow X$ and a free and proper action Φ of a Lie group G on the total space P such that

$$\pi_{XP} \circ \Phi_g = \pi_{XP}, \quad \text{for all } g \in G. \quad (7.22)$$

This is the subject of the paper [10].

8. Outlook for further studies

This paper formulated the problem of strand dynamics for an arbitrary long-range intermolecular potential in the convective representation [22] of exact geometric rod theory [37]. Its methods would also apply in the consideration of Lennard–Jones potentials and the constrained motion of non-self-interacting curves.

The paper demonstrated and compared three different approaches to deriving the same continuum equations of motion for an elastic strand experiencing non-local (for example, electrostatic or Lennard–Jones) interactions. These were: (1) the Euler–Poincaré approach, (2) the affine transformation approach, and (3) the covariant Lagrange–Poincaré formulation. In Appendix B, the Hamilton–Pontryagin approach for deriving these equations is also discussed.

It would be important to understand, from a geometric point of view, why the classical Lagrangian reduction (2) and the covariant Lagrangian reduction (3) yield the same equations of motion, modulo a change of variables. Such a question is raised and answered in [13] for the simpler case when, on the covariant side, the symmetry group coincides with the structure group of the principal bundle. We defer to future work the extension of this theory for the strand.

The present paper concentrated primarily on the case in which the strand is one-dimensional, which is a major object of interest for biological applications. However, these approaches for deriving continuum motion equations possess more significance and applicability than might be suggested by the one-dimensional developments illustrated here. For example, the geometrical considerations and nonlinear context of the present investigation would also apply in formulating the dynamics of the higher dimensional case. That is, when s has more than one component, the approaches discussed here still apply.

A change in dimensionality of s in equations (3.23) and (3.26) requires summing over all components of s -derivatives (instead of only the single s -derivative for the strand). Additional integrability conditions arise from the equality of cross-derivatives with respect to space and time that generalize equations (2.18) and (2.19). (In geometric terms, these are zero curvature conditions.) The extension to higher dimensions was discussed in the general setting treated in Section 2.4. The higher dimensional options also figured in the covariant Lagrange–Poincaré formulas (D.1), where div_x denotes derivative with respect to time and *all* dimensions of the space (taken to be one-dimensional in the paper). The extension to higher dimensions illuminates the geometry underlying the present one-dimensional case and may be expected to produce interesting applications in the dynamical description of biological membranes and other extended physical objects. While the equations take the same geometrical form in higher dimensions, their solutions will possess their own unique features.

Besides passing to higher dimensions, future studies will consider both linear and nonlinear wave propagation on electrostatically charged strands, as well as the description of nontrivial stationary states that arise from nonlocal interactions, such as for the VDF oligomers mentioned in the Introduction.

Yet another interesting question for future studies concerns the possibility of enhancing the internal structure of the rigid charge conformations. This will allow even richer dynamics than we considered here. While the resulting equations may be different (and more complex), the methods developed in this paper will still be applicable when the dynamics takes place in spaces that possess richer conformational structure than rigid rotations.

Many interesting and nontrivial issues for future research are raised by the symmetry reduced formulation of convective dynamics introduced here for nonlocal interactions of charged strands. As mentioned earlier, these issues include classification and stability analysis of equilibrium solutions, dynamics of conformational changes (folding/unfolding) and adaptation of these methods to computational approaches, all of which we must place beyond the scope of the present work.

Acknowledgements. DDH and VP were partially supported by NSF grants NSF-DMS-05377891 and NSF-DMS-0908755. The work of DDH was also partially supported by the Royal Society of London Wolfson Research Merit Award. VP acknowledges the support of the European Science Foundation for partial support through the MISGAM program. FGB acknowledges the partial support of Swiss National Science Foundation grants 200020-117511 and of a Swiss National Science Foundation Fellowship. TSR acknowledges the partial support of Swiss National Science Foundation grants 200020-117511 and 200020-126630. We thank the referees for their helpful, thoughtful remarks and suggestions.

Appendix A. List of notations

This Appendix contains a list of the key notations used in the paper. Many of the relationships among the variables and the spaces in which they live are also diagrammed in Section 7.2.3.

t	Time
s	Coordinate along a strand
$\mathbf{r}(s, t) \in \mathbb{R}^3$	Spatial position of a strand
$SO(3)$	Special orthogonal group of \mathbb{R}^3
$\mathcal{O} \ltimes E$	Semidirect product of a Lie group \mathcal{O} with a vector space E
$SE(3) \simeq SO(3) \ltimes \mathbb{R}^3$	Special Euclidean group
RCC	Rigid charge conformation
$\Lambda(s, t) \in SO(3)$	Rotation of the RCC
$\xi(s, s') = \Lambda^{-1}(s)\Lambda(s')$	Relative rotation of RCCs at s and s'
$\kappa(s, s') = \Lambda^{-1}(s)(\mathbf{r}(s') - \mathbf{r}(s))$	Chord connecting s and s' as seen in RCC frame at s
$\mathcal{F}(I, SE(3))$	$SE(3)$ -valued smooth functions on interval I
(Λ, \mathbf{r})	A curve in the group $\mathcal{F}(I, SE(3))$
G	Lie group
TG	Tangent bundle of a Lie group G
T^*G	Cotangent bundle of a Lie group G
\mathfrak{g}	Lie algebra of G
\mathfrak{g}^*	Dual of the Lie algebra of G
Ad and Ad^*	Adjoint and coadjoint representations of a Lie group
ad and ad^*	Adjoint and coadjoint representations of a Lie algebra
$(\boldsymbol{\omega}, \boldsymbol{\gamma})$	Convective angular and linear velocity distributions
$(\boldsymbol{\Omega}, \boldsymbol{\Gamma})$	Deformation gradients in convective representation
ρ	Position of the filament as viewed from the RCC
$(\boldsymbol{\mu}, \boldsymbol{\pi}) = \left(\frac{\delta l}{\delta \boldsymbol{\omega}}, \frac{\delta l}{\delta \boldsymbol{\gamma}} \right)$	Convective angular and linear momentum densities
$\{f, g\}(\boldsymbol{\mu}, \boldsymbol{\pi}, \boldsymbol{\Omega}, \boldsymbol{\Gamma}, \rho)$	Affine Poisson bracket on $(\mathcal{F}(I, \mathfrak{se}(3)) \ltimes \mathcal{F}(I, \mathfrak{se}(3) \times \mathbb{R}^3))^*$
$\text{Ad}^*_{(\Lambda, \mathbf{r})^{-1}} \left(\frac{\delta l}{\delta \boldsymbol{\omega}}, \frac{\delta l}{\delta \boldsymbol{\gamma}} \right)$	Spatial angular and linear momentum density

Appendix B. A modified Hamilton–Pontryagin approach

A second derivation of the key formulas (3.23) and (3.26) for strand dynamics in the convective representation is given in this Appendix. This derivation is based on the Hamilton–Pontryagin (HP) approach in control theory (see, for example, [5]), which will be modified to include additional terms describing the nonlocal contributions. The elegance and directness of the HP approach in this Appendix relative to the Euler–Poincaré approach in Section 3 is accomplished by simplifying the interplay between the group action and the variational principle at the expense of introducing extra variables. The equivalent Euler–Poincaré derivation is more elaborate than the HP derivation, because the Euler–Poincaré approach invokes the Lie group action on the configuration space and thereby provides additional information. Section 5 has shown that the derivation of the Euler–Poincaré equations and of the associated variational principle are corollaries of general theorems for systems whose configuration space is a Lie group. The complementary, but less transparent, HP route in that case reveals other perspectives and results whose abstract general formulation will be explored in future work. Some calculations in this Appendix overlap with those in Section 3. Nonetheless, we have chosen to present them here for completeness and ease of exposition.

Appendix B.1. Filament dynamics

We first apply the Hamilton–Pontryagin approach to the case when the Lagrangian includes only the local part, so that $l = l_{\text{loc}}(\omega, \gamma, \Omega, \Gamma, \rho)$. Inspired by the classical Hamilton–Pontryagin approach, we introduce Lagrange multipliers for the holonomic constraints that impose the defining relations (2.11) for the five reduced variables $(\omega, \gamma, \Omega, \Gamma, \rho)$ in equations (1.1).

Theorem B.1. (Hamilton–Pontryagin theorem for filament dynamics) *The equations for filament dynamics arise from the variational principle $\delta S = 0$ with action S given by*

$$S = \int l(\omega, \gamma, \Omega, \Gamma, \rho) dt + \iint \left(\pi \cdot (\Lambda^{-1} \dot{\Lambda} - \omega) + \Pi \cdot (\Lambda^{-1} \Lambda' - \Omega) \right. \\ \left. + \mathbf{R} \cdot (\Lambda^{-1} \mathbf{r} - \rho) + \mu \cdot (\Lambda^{-1} \dot{\mathbf{r}} - \gamma) + \mathbf{M} \cdot (\Lambda^{-1} \mathbf{r}' - \Gamma) \right) ds dt.$$

These equations are

$$\frac{\delta l}{\delta \rho} - \mathbf{R} = 0, \quad \frac{\delta l}{\delta \omega} - \pi = 0, \quad \frac{\delta l}{\delta \Omega} - \Pi = 0, \quad \frac{\delta l}{\delta \gamma} - \mu = 0, \quad \frac{\delta l}{\delta \Gamma} - \mathbf{M} = 0, \\ \dot{\pi} + \omega \times \pi + \Pi' + \Omega \times \Pi + \gamma \times \mu + \Gamma \times \mathbf{M} + \rho \times \mathbf{R} = 0,$$

and

$$\dot{\mu} + \omega \times \mu + \mathbf{M}' + \Omega \times \mathbf{M} - \mathbf{R} = 0.$$

together with the constraints,

$$\Lambda^{-1} \dot{\Lambda} = \omega, \quad \Lambda^{-1} \Lambda' = \Omega, \quad \Lambda^{-1} \mathbf{r} = \rho, \quad \Lambda^{-1} \dot{\mathbf{r}} = \gamma, \quad \Lambda^{-1} \mathbf{r}' = \Gamma.$$

To facilitate the proof of the theorem, we first prove a lemma that will be helpful in computing the variations of the quantities appearing in the action S .

Lemma B.1. *The variations of the quantities in Λ and \mathbf{r} of the formulas in (2.11) are*

$$\begin{aligned}\delta\left(\Lambda^{-1}\dot{\Lambda}\right) &= \frac{\partial\widehat{\Sigma}}{\partial t} + \left[\Lambda^{-1}\dot{\Lambda}, \widehat{\Sigma}\right], \\ \delta\left(\Lambda^{-1}\Lambda'\right) &= \widehat{\Sigma}' + \left[\Lambda^{-1}\Lambda', \widehat{\Sigma}\right], \\ \delta\left(\Lambda^{-1}\mathbf{r}\right) &= \Psi - \widehat{\Sigma}\left(\Lambda^{-1}\mathbf{r}\right), \\ \delta\left(\Lambda^{-1}\dot{\mathbf{r}}\right) &= \dot{\Psi} - \widehat{\Sigma}\left(\Lambda^{-1}\dot{\mathbf{r}}\right) + \left(\Lambda^{-1}\dot{\Lambda}\right)\Psi, \\ \delta\left(\Lambda^{-1}\mathbf{r}'\right) &= \Psi' - \widehat{\Sigma}\left(\Lambda^{-1}\mathbf{r}'\right) + \left(\Lambda^{-1}\Lambda'\right)\Psi.\end{aligned}\tag{B.1}$$

Proof. We calculate the variations directly, one by one. First we have,

$$\begin{aligned}\delta\left(\Lambda^{-1}\dot{\Lambda}\right) &= -\Lambda^{-1}\delta\Lambda\left(\Lambda^{-1}\dot{\Lambda}\right) + \Lambda^{-1}\delta\dot{\Lambda} \\ &= -\Lambda^{-1}\delta\Lambda\left(\Lambda^{-1}\dot{\Lambda}\right) + \left(\Lambda^{-1}\delta\Lambda\right)' + \left(\Lambda^{-1}\dot{\Lambda}\right)\left(\Lambda^{-1}\delta\Lambda\right) \\ &= \frac{\partial\widehat{\Sigma}}{\partial t} + \left[\Lambda^{-1}\dot{\Lambda}, \widehat{\Sigma}\right].\end{aligned}$$

Similarly, for the variation of $\Lambda^{-1}\Lambda'$ we have,

$$\delta\left(\Lambda^{-1}\Lambda'\right) = \widehat{\Sigma}' + \left[\Lambda^{-1}\Lambda', \widehat{\Sigma}\right].$$

Now we consider the variation of $\Lambda^{-1}\dot{\mathbf{r}}$, which is given by

$$\begin{aligned}\delta\left(\Lambda^{-1}\dot{\mathbf{r}}\right) &= -\Lambda^{-1}\delta\Lambda\Lambda^{-1}\mathbf{r} + \Lambda^{-1}\delta\dot{\mathbf{r}} \\ &= -\left(\Lambda^{-1}\delta\Lambda\right)\left(\Lambda^{-1}\mathbf{r}\right) + \left(\Lambda^{-1}\delta\mathbf{r}\right)' + \left(\Lambda^{-1}\dot{\Lambda}\right)\left(\Lambda^{-1}\mathbf{r}\right) \\ &= \dot{\Psi} - \widehat{\Sigma}\left(\Lambda^{-1}\dot{\mathbf{r}}\right) + \left(\Lambda^{-1}\dot{\Lambda}\right)\Psi.\end{aligned}$$

A similar argument yields the variation of $\Lambda^{-1}\mathbf{r}'$,

$$\delta\left(\Lambda^{-1}\mathbf{r}'\right) = \Psi' - \widehat{\Sigma}\left(\Lambda^{-1}\mathbf{r}'\right) + \left(\Lambda^{-1}\Lambda'\right)\Psi.$$

Finally, the variation of $\Lambda^{-1}\mathbf{r}$ is given by,

$$\begin{aligned}\delta\left(\Lambda^{-1}\mathbf{r}\right) &= -\Lambda^{-1}\delta\Lambda\Lambda^{-1}\mathbf{r} + \Lambda^{-1}\delta\mathbf{r} \\ &= \Psi - \widehat{\Sigma}\left(\Lambda^{-1}\mathbf{r}\right)\end{aligned}$$

and all the formulas in the statement are proved. \square

We may now use these variational formulas in (B.1) to prove the Hamilton–Pontryagin Theorem B.1 for the equations of filament dynamics.

Proof. Variations with respect to the Lagrange multipliers impose the defining relations for the five quantities $\{\boldsymbol{\rho}, \boldsymbol{\omega}, \boldsymbol{\Omega}, \boldsymbol{\gamma}, \boldsymbol{\Gamma}\}$. The conjugate variations give

$$\frac{\delta l}{\delta \boldsymbol{\rho}} - \mathbf{R} = 0, \quad \frac{\delta l}{\delta \boldsymbol{\omega}} - \boldsymbol{\pi} = 0, \quad \frac{\delta l}{\delta \boldsymbol{\Omega}} - \boldsymbol{\Pi} = 0, \quad \frac{\delta l}{\delta \boldsymbol{\gamma}} - \boldsymbol{\mu} = 0, \quad \frac{\delta l}{\delta \boldsymbol{\Gamma}} - \mathbf{M} = 0.$$

Collecting the variations proportional to $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$ yields the filament equations

$$\dot{\boldsymbol{\pi}} + \boldsymbol{\omega} \times \boldsymbol{\pi} + \boldsymbol{\Pi}' + \boldsymbol{\Omega} \times \boldsymbol{\Pi} + \boldsymbol{\gamma} \times \boldsymbol{\mu} + \boldsymbol{\Gamma} \times \mathbf{M} + \boldsymbol{\rho} \times \mathbf{R} = 0,$$

and

$$\dot{\boldsymbol{\mu}} + \boldsymbol{\omega} \times \boldsymbol{\mu} + \mathbf{M}' + \boldsymbol{\Omega} \times \mathbf{M} - \mathbf{R} = 0,$$

respectively. \square

Remark B.1. The Hamilton–Pontryagin approach used here also allows nonholonomic constraints to be imposed on the motion of the strand, although we shall refrain from pursuing that direction here. See [19] for a discussion of nonholonomic constraints using the Hamilton–Pontryagin approach.

Appendix B.2. Nonlocal potential

For the nonlocal potential (2.13) we may form a Hamilton–Pontryagin variational principle in a similar fashion. In this case, the action S_{np} is given by

$$\begin{aligned} S_{\text{np}} = & \int l_{\text{np}}(\xi, \boldsymbol{\kappa}, \boldsymbol{\Gamma}) dt + \iint \mathbf{m} \cdot \left(\Lambda^{-1}(s) \mathbf{r}'(s) - \boldsymbol{\Gamma} \right) ds dt \\ & + \iiint \left(\mathbf{X} \cdot \left(\Lambda^{-1}(s) \Lambda(s') - \xi \right) + \mathbf{K} \cdot \left(\Lambda^{-1}(s) (\mathbf{r}(s') - \mathbf{r}(s)) - \boldsymbol{\kappa} \right) \right) ds ds' dt. \end{aligned}$$

Since the strand is not assumed to be locally inextensible, the stretch of the strand's base requires extra factors of $|\boldsymbol{\Gamma}|$ multiplying the differential of the parameter $|\boldsymbol{\Gamma}(s)|$ along the strand ds in the expression for potential energy. However, in this section, in order to simplify the formulas and avoid extra factors in the integrals, we shall incorporate that factor of $|\boldsymbol{\Gamma}(s)|$ into the nonlocal potential. No confusion with the previous sections should arise here.

Lemma B.2. *The additional variational formulas needed for calculating the equations of motion are given by*

$$\begin{aligned} \Lambda^{-1}(s') \Lambda(s) \left(\delta \left(\Lambda^{-1}(s) \Lambda(s') \right) \right) &= -\text{Ad}_{\Lambda^{-1}(s') \Lambda(s)} \widehat{\boldsymbol{\Sigma}}(s) + \widehat{\boldsymbol{\Sigma}}(s'), \\ \delta \left(\Lambda^{-1}(s) (\mathbf{r}(s') - \mathbf{r}(s)) \right) &= -\widehat{\boldsymbol{\Sigma}}(s) \Lambda^{-1}(s) (\mathbf{r}(s') - \mathbf{r}(s)) \\ &\quad + \Lambda^{-1}(s) \Lambda(s') \boldsymbol{\Psi}(s') + \boldsymbol{\Psi}(s). \end{aligned}$$

where the independent variations are defined by

$$\boldsymbol{\Psi}(s) = \Lambda^{-1}(s) \delta \mathbf{r}(s) \quad \text{and} \quad \widehat{\boldsymbol{\Sigma}}(s) = \Lambda^{-1}(s) \delta \Lambda(s). \quad (\text{B.2})$$

Proof. The first variational formula is calculated directly, as

$$\begin{aligned}\Lambda^{-1}(s')\Lambda(s) \left(\delta \left(\Lambda^{-1}(s)\Lambda(s') \right) \right) &= \Lambda^{-1}(s')\Lambda(s) \left(\Lambda^{-1}(s)\delta\Lambda(s') \right) \\ &\quad - \Lambda^{-1}(s')\Lambda(s) \left(\Lambda^{-1}(s)\delta\Lambda(s)\Lambda^{-1}(s)\Lambda(s') \right) \\ &= -\text{Ad}_{\Lambda^{-1}(s')\Lambda(s)} \widehat{\Sigma}(s) + \widehat{\Sigma}(s').\end{aligned}$$

The second variational formula follows similarly from a direct calculation, namely,

$$\begin{aligned}\delta \left(\Lambda^{-1}(s) \left(\mathbf{r}(s') - \mathbf{r}(s) \right) \right) &= -\Lambda^{-1}(s)\delta\Lambda(s)\Lambda^{-1}(s) \left(\mathbf{r}(s') - \mathbf{r}(s) \right) \\ &\quad + \Lambda^{-1}(s) \left(\delta\mathbf{r}(s') - \delta\mathbf{r}(s) \right) \\ &= -\widehat{\Sigma}(s)\Lambda^{-1}(s) \left(\mathbf{r}(s') - \mathbf{r}(s) \right) \\ &\quad + \Lambda^{-1}(s)\Lambda(s')\Psi(s') + \Psi(s)\end{aligned}$$

which proves the lemma. \square

Theorem B.2. *The equations that arise from the variational principle with the nonlocal action*

$$\begin{aligned}S_{\text{np}} &= \iint U(\xi(s, s'), \kappa(s, s'), \Gamma(s)) ds ds' dt + \iint \mathbf{m} \cdot \left(\Lambda^{-1}(s)\mathbf{r}'(s) - \Gamma(s) \right) ds dt \\ &\quad + \iint \left(X \cdot \left(\Lambda^{-1}(s)\Lambda(s') - \xi(s, s') \right) \right. \\ &\quad \left. + \mathbf{K} \cdot \left(\Lambda^{-1}(s) \left(\mathbf{r}(s') - \mathbf{r}(s) \right) - \kappa(s, s') \right) \right) ds ds' dt\end{aligned}$$

are given by:

$$X = \frac{\partial U}{\partial \xi}, \quad \mathbf{K} = \frac{\partial U}{\partial \kappa}, \quad \mathbf{m} = \frac{\partial U}{\partial \Gamma},$$

and

$$\begin{aligned}\Gamma \times \mathbf{m} &= \int \left(\xi(s, s')X(s', s) - X(s, s')\xi^{-1}(s, s') + \mathbf{K}(s, s') \times \kappa(s, s') \right) ds', \\ \mathbf{m}' + \Omega \times \mathbf{m} &= \int \left(\xi(s, s')\mathbf{K}(s', s) - \mathbf{K}(s, s') \right) ds',\end{aligned}$$

together with the constraints,

$$\xi(s, s') = \Lambda^{-1}(s)\Lambda(s'), \quad \kappa(s, s') = \Lambda^{-1}(s) \left(\mathbf{r}(s') - \mathbf{r}(s) \right), \quad \Gamma(s) = \Lambda^{-1}(s)\mathbf{r}'(s).$$

Proof. The proof is obtained by substituting the variations given in Lemma B.2 into Hamilton's principle for the action in the statement of the theorem. Variations in X , \mathbf{K} , and \mathbf{m} yield the constraints

$$\xi(s, s') = \Lambda^{-1}(s)\Lambda(s'), \quad \kappa(s, s') = \Lambda^{-1}(s) \left(\mathbf{r}(s') - \mathbf{r}(s) \right), \quad \Gamma(s) = \Lambda^{-1}(s)\mathbf{r}'(s).$$

Variations in ξ , κ , and Γ yield the relationships

$$X = \frac{\partial U}{\partial \xi}, \quad K = \frac{\partial U}{\partial \kappa}, \quad m = \frac{\partial U}{\partial \Gamma}.$$

Finally, the variations proportional to $\widehat{\Sigma}(s)$ and $\Psi(s)$ yield

$$\Gamma \times m = \int \left(\xi(s, s') X(s', s) - X(s, s') \xi^{-1}(s, s') + K(s, s') \times \kappa(s, s') \right) ds'$$

and

$$m' + \Omega \times m = \int \left(\xi(s, s') K(s', s) - K(s, s') \right) ds',$$

respectively. \square

We may combine these nonlocal terms and the local part of the equations to produce the full set of equations. These are given by

$$\begin{aligned} \dot{\pi} + \omega \times \pi + \Pi' + \Omega \times \Pi + \gamma \times \mu + \Gamma \times (\mathbf{M} + m) + \rho \times \mathbf{R} \\ = \int \left(K(s, s') \times \kappa(s, s') + \mathbf{Z}(s, s') \right) ds' \end{aligned}$$

and

$$\begin{aligned} \dot{\mu} + \omega \times \mu + (\mathbf{M} + m)' + \Omega \times (\mathbf{M} + m) - \mathbf{R} \\ = \int \left(\xi(s, s') K(s', s) - K(s, s') \right) ds', \end{aligned}$$

where one defines, as in formula (3.24),

$$\widehat{\mathbf{Z}}(s, s') := \xi(s, s') X(s', s) - X(s, s') \xi^{-1}(s, s'), \quad (\text{B.3})$$

denoted as $\widehat{\mathbf{Z}}$ since the right-hand side of this equation is in $\mathfrak{so}(3)$.

We may now use these functional derivative relations to express the equations of motion in terms of the reduced Lagrangian, $l = l_{\text{loc}} + l_{\text{np}}$. The functional derivative relations obtained in the Hamilton–Pontryagin approach are

$$\begin{aligned} \mathbf{R} &= \frac{\delta l_{\text{loc}}}{\delta \rho}, & \pi &= \frac{\delta l_{\text{loc}}}{\delta \omega}, \\ \Pi &= \frac{\delta l_{\text{loc}}}{\delta \Omega}, & \mu &= \frac{\delta l_{\text{loc}}}{\delta \gamma}, \\ \mathbf{M} &= \frac{\delta l_{\text{loc}}}{\delta \Gamma}, & X &= \frac{\delta l_{\text{np}}}{\delta \xi}, \\ K &= \frac{\delta l_{\text{np}}}{\delta \kappa}, & \mathbf{M} + m &= \frac{\delta (l_{\text{loc}} + l_{\text{np}})}{\delta \Gamma}. \end{aligned}$$

Substituting these relations into the equations of motion above gives the following equations of motion for the charged strand.

$$(\partial_t + \boldsymbol{\omega} \times) \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\omega}} + (\partial_s + \boldsymbol{\Omega} \times) \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\Omega}} = \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\gamma}} \times \boldsymbol{\gamma} + \frac{\delta (l_{\text{loc}} + l_{\text{np}})}{\delta \boldsymbol{\Gamma}} \times \boldsymbol{\Gamma} + \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\rho}} \times \boldsymbol{\rho} \\ + \int \left(\frac{\partial U}{\partial \boldsymbol{\kappa}}(s, s') \times \boldsymbol{\kappa}(s, s') + \mathbf{Z}(s, s') \right) ds', \quad (\text{B.4})$$

$$(\partial_t + \boldsymbol{\omega} \times) \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\gamma}} + (\partial_s + \boldsymbol{\Omega} \times) \frac{\delta (l_{\text{loc}} + l_{\text{np}})}{\delta \boldsymbol{\Gamma}} \\ = \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\rho}} + \int \left(\xi(s, s') \frac{\partial U}{\partial \boldsymbol{\kappa}}(s', s) - \frac{\partial U}{\partial \boldsymbol{\kappa}}(s, s') \right) ds'. \quad (\text{B.5})$$

The term $\widehat{\mathbf{Z}}(s, s')$ is the contribution from the nonlocal part of the Lagrangian that we have sought.

The dynamical equations (B.4) and (B.5) obtained by the Hamilton–Pontryagin approach recover equations (3.23) and (3.26), respectively, from the Euler–Poincaré approach. These equations must be augmented by the advection conditions (2.17), (2.18), and (2.19) in order to close the system.

Appendix C. Formulation of nonlocal exact geometric rods in terms of quaternions

Quaternions allow for a simple, elegant, and useful method of describing the local orientation of a curve. It is thus natural to seek a representation of our derivation in previous sections that expresses the strand equations in terms of quaternions. The quaternion representation is natural, for example, in formulating the equations of motion for elastic rods in terms of the corresponding Euler parameters. As far as we are aware, a treatment of continuum rod theory in terms of quaternions in the nonlocal sense presented here does not appear in the literature. We shall see how the nonlocal contribution (3.24) appears as an imaginary part of a certain quaternion, thereby making the connection to other work. This is accomplished by mapping quaternions (elements of $SU(2)$) that describe rotations into purely imaginary quaternions, or vectors, that are elements of $\mathfrak{su}(2) \simeq \mathfrak{so}(3) \simeq \mathbb{R}^3$.

Remark C.1. This section simplifies the formulas and avoids extra factors in the integrals by assuming the strand to be inextensible; thus $|\boldsymbol{\Gamma}(s)|$ is identically equal to one and the parameter s is the arc length. See (3.23) and (3.26) for the case $|\boldsymbol{\Gamma}(s)| \neq 1$.

Let us associate a quaternion $\mathbf{q} = (q_0, \mathbf{q})$ with every point on the curve s . That quaternion describes the local rotation of an orthogonal frame if the condition $\|\mathbf{q}\| = q_0^2 + |\mathbf{q}|^2 = 1$ is satisfied. Then, $q_0 = \cos(\alpha/2)$, with rotation angle α and $\mathbf{q} = \sin(\alpha/2)\hat{\mathbf{n}}$, where $\hat{\mathbf{n}}$ is a unit vector which is the axis of rotation. These are the Cayley–Klein parameters of the rotation.

Remark C.2. To simplify the notation, we use bold symbols for purely imaginary quaternions, considering them as vectors. For example, if q is a unit quaternion, then

$$\mathbf{b} = qaq^*$$

means

$$(0, \mathbf{b}) = q(0, \mathbf{a})q^*,$$

where $(0, \mathbf{a})$ and $(0, \mathbf{b})$ are two unit imaginary quaternions.

As before, we assume that the interaction potential depends on the distances between point charges that are attached to each point $\mathbf{r}(s, t)$ by rigid rods of the length $\eta_i(s)$. The new position of the charges will be $q\eta_i q^*$, where concatenation denotes quaternion multiplication and q^* is the quaternionic conjugate of q . The point charges are then positioned at the coordinates in real space $\mathbf{r}(s, t) + q\eta_i q^*$ and the distance between point charges is then³

$$d_{k,m}(s, s') = |\mathbf{r}(s) - \mathbf{r}(s') + q(s)\eta_k(s)q^*(s) - q(s')\eta_m(s')q^*(s')|. \quad (\text{C.1})$$

This is simply (2.3) written now in its quaternionic form. Following (2.6), we perform Lie–Poisson reduction as follows (remember that $\|q\| = 1$ and $qq^* = \epsilon$, where $\epsilon = (1, 0)$ is the unit quaternion):

$$\begin{aligned} d_{k,m}(s, s') &= |\mathbf{r}(s') - \mathbf{r}(s) + q(s')\eta_k(s')q^*(s') - q(s)\eta_m(s)q^*(s)| \\ &= |q^*(s)(0, \mathbf{r}(s') - \mathbf{r}(s) + q(s')\eta_k(s')q^*(s') - q(s)\eta_m(s)q^*(s))q(s)| \\ &= |\mathfrak{z}(s, s')\boldsymbol{\rho}(s')\mathfrak{z}^*(s, s') - \boldsymbol{\rho}(s) + \mathfrak{z}(s, s')\eta_k(s')\mathfrak{z}^*(s, s') - \eta_m(s)| \\ &= |\boldsymbol{\kappa}(s, s') + \eta_k(s) - \mathfrak{z}(s, s')\eta_m(s')\mathfrak{z}^*(s, s')|, \end{aligned} \quad (\text{C.2})$$

where $\mathfrak{z}(s, s') = q^*(s)q(s')$ is the coupling between the frames and the quantity

$$\boldsymbol{\rho}(s) = q^*(s)\mathbf{r}(s)q(s),$$

is the distance vector connecting the points $\mathbf{r}(s)$ and $\mathbf{r}(s')$ transformed according to the inverse rotation of the frame at the point s . We have also defined

$$\boldsymbol{\kappa}(s, s') = \mathfrak{z}(s, s')\boldsymbol{\rho}(s')\mathfrak{z}^*(s, s') - \boldsymbol{\rho}(s). \quad (\text{C.3})$$

The nonlocal part of the reduced Lagrangian depends on the variables $\boldsymbol{\rho}$ and \mathfrak{z} . The local part which describes elastic deformation and inertia, can be reduced to functions of $\mathfrak{m} := q^*q'$ and $\mathfrak{v} := q^*\dot{q}$, where the prime denotes the derivative with respect to s and the dot is the derivative with respect to t . The quaternions \mathfrak{m} and \mathfrak{v} belong to the Lie algebra of the Lie group of all unit quaternions which is isomorphic to the space of purely imaginary quaternions, or vectors, as can be seen by differentiating $q^*q = \epsilon$. The commutator is then mapped into twice the vector product of the imaginary parts of the quaternions.

³ All these variables depend on time t as well as s , but the time variable t is suppressed.

We can again split the reduced Lagrangian l in the local and nonlocal parts

$$l = \int l_{\text{loc}}(\mathbf{v}(s), \mathbf{m}(s), \boldsymbol{\gamma}(s), \boldsymbol{\Gamma}(s), \boldsymbol{\rho}(s)) \, ds + \iint U(\boldsymbol{\kappa}(s, s'), \mathbf{z}(s, s')) \, ds \, ds' := l_{\text{loc}} + l_{\text{np}}. \quad (\text{C.4})$$

Here, $\boldsymbol{\gamma}$, $\boldsymbol{\Gamma}$, $\boldsymbol{\rho}$ are defined as in (2.11). The equations of motions then follow from the minimization of the reduced action

$$\delta S = \delta \int l(\mathbf{v}(s), \mathbf{m}(s), \boldsymbol{\rho}(s), \boldsymbol{\kappa}(s, s'), \mathbf{z}(s, s')) \, ds \, ds' \, dt = 0. \quad (\text{C.5})$$

Here \mathbf{v} and \mathbf{s} are elements of Lie algebra of purely imaginary quaternions with absolute value equal to one. They are purely imaginary quaternions, $\mathbf{v} = (0, \boldsymbol{\omega}/2)$ and $\mathbf{m} = (0, \boldsymbol{\Omega}/2)$; the real Lie algebra of imaginary quaternions is Lie algebra isomorphic as to (\mathbb{R}^3, \times) . The factor $1/2$ is necessary for $\boldsymbol{\omega}$ and $\boldsymbol{\Omega}$ to be exactly the vector angular velocity and strain rate, respectively, in agreement with our notation. Thus, we can write (C.5) using vector quantities instead of quaternions whenever possible:

$$\delta S = \delta \iint l_{\text{np}}(\boldsymbol{\omega}, \boldsymbol{\gamma}, \boldsymbol{\Omega}, \boldsymbol{\Gamma}, \boldsymbol{\rho}) \, ds \, dt + \delta \iiint U(\boldsymbol{\kappa}(s, s'), \mathbf{z}(s, s')) \, ds \, ds' \, dt = 0. \quad (\text{C.6})$$

We obtain

$$\begin{aligned} \delta S = & \iint \left\langle \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\Omega}}, \delta \boldsymbol{\Omega} \right\rangle + \left\langle \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\omega}}, \delta \boldsymbol{\omega} \right\rangle + \left\langle \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\rho}}, \delta \boldsymbol{\rho} \right\rangle + \left\langle \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\gamma}}, \delta \boldsymbol{\gamma} \right\rangle \\ & + \left\langle \frac{\delta l_{\text{loc}}}{\delta \boldsymbol{\Gamma}}, \delta \boldsymbol{\Gamma} \right\rangle \, ds \, dt + \iiint \left\langle \frac{\delta l_{\text{np}}}{\delta \mathbf{z}}, \delta \mathbf{z} \right\rangle + \left\langle \frac{\delta l_{\text{np}}}{\delta \boldsymbol{\kappa}}, \delta \boldsymbol{\kappa} \right\rangle \, ds \, ds' \, dt = 0. \end{aligned} \quad (\text{C.7})$$

If we now define $\mathbf{s} := \mathbf{q}^* \delta \mathbf{q} = \mathbf{q}^{-1} \delta \mathbf{q}$ as the free variation in \mathbf{q} (a unit quaternion), we obtain, similarly to Section 3.1.1:

$$\delta \mathbf{v} = \mathbf{v} \mathbf{s} - \mathbf{s} \mathbf{v} + \dot{\mathbf{s}} = [\mathbf{v}, \mathbf{s}] + \dot{\mathbf{s}}, \quad (\text{C.8})$$

for the time derivative and

$$\delta \mathbf{m} = \mathbf{m} \mathbf{s} - \mathbf{s} \mathbf{m} + \mathbf{s}' = [\mathbf{m}, \mathbf{s}] + \mathbf{s}', \quad (\text{C.9})$$

for the space derivative. Note that since $\mathbf{q}^* \mathbf{q} = 1$,

$$\mathbf{q}^* \dot{\mathbf{q}} + \dot{\mathbf{q}}^* \mathbf{q} = \mathbf{q}^* \dot{\mathbf{q}} + (\mathbf{q}^* \dot{\mathbf{q}})^* = 2 \text{Re } \mathbf{v} = 0$$

and, analogously, $\text{Re } \mathbf{m} = 0$, which means that \mathbf{v} and \mathbf{m} are purely imaginary quaternions, or vectors. This allows us to compute the first two terms in (C.7). We now remember that $\text{Re } \mathbf{v} = 0$ and $\text{Re } \mathbf{s} = 0$ since they are elements of the corresponding

Lie algebra, so $\mathfrak{v} = (0, \boldsymbol{\omega}/2)$, $\mathfrak{m} = (0, \boldsymbol{\Omega}/2)$, and $\mathfrak{s} = (0, \boldsymbol{\Sigma}/2)$. Then, (C.8) and (C.9) can be expressed as vector equations:

$$\delta \mathfrak{v} = (0, \delta \boldsymbol{\omega}) = (0, \boldsymbol{\omega} \times \boldsymbol{\Sigma} + \dot{\boldsymbol{\Sigma}}), \quad (\text{C.10})$$

and

$$\delta \mathfrak{m} = (0, \delta \boldsymbol{\Omega}) = (0, \boldsymbol{\Omega} \times \boldsymbol{\Sigma} + \boldsymbol{\Sigma}'). \quad (\text{C.11})$$

Now, computations of the first three variations in (C.7) can be done analogously to those in Section 3.1.1, as they involve vector quantities. The only exception is the computation of $\delta \boldsymbol{\rho}$ in the nonlocal term as it must be computed in terms of quaternions. We have

$$\begin{aligned} \delta \boldsymbol{\rho} &= \delta (\mathfrak{q}^*(s) \mathbf{r}(s) \mathfrak{q}) \\ &= -\mathfrak{q}^* \delta \mathfrak{q} \boldsymbol{\rho}(s) + \boldsymbol{\rho}(s) \mathfrak{q}^* \delta \mathfrak{q} + \mathfrak{q}^*(s) \delta \mathbf{r}(s) \mathfrak{q}(s) = 2\boldsymbol{\rho} \times \boldsymbol{\Sigma} + \boldsymbol{\Psi}, \end{aligned} \quad (\text{C.12})$$

where we have defined the free variation

$$\boldsymbol{\Psi}(s) := \mathfrak{q}^*(s) \delta \mathbf{r}(s) \mathfrak{q}(s). \quad (\text{C.13})$$

We have also used the relation $\mathfrak{s} \boldsymbol{\rho} - \boldsymbol{\rho} \mathfrak{s} = \boldsymbol{\rho} \times \boldsymbol{\Sigma}$ that is satisfied for purely imaginary $\mathfrak{s} = (0, \boldsymbol{\Sigma})$. Next, we need to compute the variation of the nonlocal part of $\boldsymbol{\kappa}(s, s')$ as follows. It is easier to use the alternative expression for $\boldsymbol{\kappa}$ as

$$\boldsymbol{\kappa}(s, s') = \boldsymbol{\rho}(s, s') - \mathfrak{q}^*(s) \mathbf{r}(s') \mathfrak{q}(s).$$

Then,

$$\begin{aligned} \delta \boldsymbol{\kappa} &= \delta \boldsymbol{\rho} - \delta (\mathfrak{q}^*(s) \mathbf{r}(s') \mathfrak{q}(s)) \\ &= \boldsymbol{\rho}(s) \times \boldsymbol{\Sigma}(s) + \boldsymbol{\Psi}(s) - \delta (\mathfrak{q}^*(s, s') \boldsymbol{\rho}(s') \mathfrak{z}(s, s')) \\ &= \boldsymbol{\rho}(s) \times \boldsymbol{\Sigma}(s) + \boldsymbol{\Psi}(s) - \delta \mathfrak{q}^*(s) \mathbf{r}(s') \mathfrak{q}(s) - \mathfrak{q}^*(s) \delta \mathbf{r}(s') \mathfrak{q}(s) - \mathfrak{q}^*(s) \mathbf{r}(s') \delta \mathfrak{q}(s) \\ &= \boldsymbol{\rho}(s) \times \boldsymbol{\Sigma}(s) + \boldsymbol{\Psi}(s) \\ &\quad - \mathfrak{s}^*(s) (\boldsymbol{\rho}(s) - \boldsymbol{\kappa}(s, s')) - (\boldsymbol{\rho}(s) - \boldsymbol{\kappa}(s, s')) \mathfrak{s}(s) - \mathfrak{z}(s, s') \boldsymbol{\Psi}(s') \mathfrak{z}^*(s, s') \\ &= \boldsymbol{\Psi}(s) - \mathfrak{z}^*(s, s') \boldsymbol{\Psi}(s') \mathfrak{z}(s, s') - \boldsymbol{\kappa}(s, s') \times \boldsymbol{\Sigma}, \end{aligned} \quad (\text{C.14})$$

which is a direct analogue of (3.12). We have used the fact that for purely imaginary quaternions \mathfrak{s} , we have

$$\mathfrak{s}^* \boldsymbol{\rho} + \boldsymbol{\rho} \mathfrak{s} = -\mathfrak{s} \boldsymbol{\rho} + \boldsymbol{\rho} \mathfrak{s} = \boldsymbol{\rho} \times \boldsymbol{\Sigma}$$

and

$$\mathfrak{s}^*(s) \boldsymbol{\kappa}(s, s') + \boldsymbol{\kappa}(s, s') \mathfrak{s} = -\mathfrak{s} \boldsymbol{\kappa}(s, s') + \boldsymbol{\kappa}(s, s') \mathfrak{s} = \boldsymbol{\kappa}(s, s') \times \boldsymbol{\Sigma}.$$

Next, let us define the purely imaginary quaternion

$$\mathfrak{z}^*(s, s') \delta \mathfrak{z}(s, s') := 2\mathbf{T}(s, s').$$

The real part of $\mathfrak{z}^*(s, s')\delta\mathfrak{z}(s, s')$ vanishes since $\mathfrak{z}(s, s')$ is a unit quaternion. The last step is the computation of \mathbf{T} in terms of free variations Σ , which proceeds as follows:

$$\begin{aligned}\mathfrak{z}^*(s, s')\delta\mathfrak{z}(s, s') &= \mathfrak{z}^*(s, s')\delta[\mathfrak{q}^*(s)\mathfrak{q}(s')] \\ &= \mathfrak{z}^*(s, s')[-\mathfrak{q}^*(s)\delta\mathfrak{q}(s)\mathfrak{z}(s, s') + \mathfrak{q}^*(s)\mathfrak{q}(s')\mathfrak{q}^*(s')\delta\mathfrak{q}(s')] \\ &= -\frac{1}{2}\mathfrak{z}^*(s, s')\Sigma(s)\mathfrak{z}(s, s') + \frac{1}{2}\Sigma(s').\end{aligned}\quad (\text{C.15})$$

Thus, we find

$$\begin{aligned}2\mathbf{T}(s, s') &:= \text{Im}(\mathfrak{z}^*(s, s')\delta\mathfrak{z}(s, s')) \\ &= -\mathfrak{z}^*(s, s')\Sigma(s)\mathfrak{z}(s, s') + \Sigma(s') \\ &=: -\text{Ad}_{\mathfrak{z}^*}\Sigma(s) + \Sigma(s').\end{aligned}$$

Note the exact correspondence between this formula and (3.10) defining the variation $\xi^{-1}\delta\xi$. Therefore, the variation with respect to $\delta\kappa$ gives

$$\begin{aligned}\left\langle \frac{\delta l_{\text{np}}}{\delta\kappa}, \delta\kappa \right\rangle &= \int \left\langle \int \frac{\partial U}{\partial\kappa}(s, s') \times \kappa(s, s') ds', \Sigma(s) \right\rangle ds \\ &\quad + \int \left\langle \int \frac{\partial U}{\partial\kappa}(s, s') - \mathfrak{z}(s, s') \frac{\partial U}{\partial\kappa}(s', s) \mathfrak{z}^*(s, s') ds', \Psi(s) \right\rangle ds.\end{aligned}\quad (\text{C.16})$$

Analogously,

$$\begin{aligned}\iint \left\langle \mathfrak{z}^* \frac{\partial U}{\partial\mathfrak{z}}, \mathfrak{z}^* \delta\mathfrak{z} \right\rangle ds ds' &= \iint \left\langle \mathfrak{z}^* \frac{\partial U}{\partial\mathfrak{z}}, 2\mathbf{T}(s, s') \right\rangle ds ds' \\ &= \iint \left\langle \mathfrak{z}^* \frac{\partial U}{\partial\mathfrak{z}}, -\mathfrak{z}^*(s, s')\Sigma(s)\mathfrak{z}(s, s') + \Sigma(s') \right\rangle ds ds' \\ &= \iint \left\langle \mathfrak{z}^*(s', s) \frac{\partial U}{\partial\mathfrak{z}}(s', s) - \frac{\partial U}{\partial\mathfrak{z}}(s, s') \mathfrak{z}^*(s, s'), \Sigma(s) \right\rangle ds ds' \\ &= -\int \left\langle \int \frac{\partial U}{\partial\mathfrak{z}}(s, s') \mathfrak{z}^*(s, s') ds', \Sigma(s) \right\rangle ds \\ &= -\int \left\langle \int \text{Im} \left[\frac{\partial U}{\partial\mathfrak{z}}(s, s') \mathfrak{z}^*(s, s') \right] ds', \Sigma(s) \right\rangle ds.\end{aligned}\quad (\text{C.17})$$

Collecting together the terms proportional to $\Sigma(s)$ and $\Psi(s)$ in the minimal action principle (C.7) gives the system (B.4), (B.5). The role of antisymmetric matrix $Z(s, s')$ describing the nonlocal interactions in (B.5) is now played by the purely imaginary quaternion

$$\mathbf{Z}(s, s') = \text{Im} \left[\frac{\partial U}{\partial\mathfrak{z}}(s, s') \mathfrak{z}^*(s, s') \right].$$

A Hamiltonian description closely following that of Section 3.3 can be developed in quaternionic form, as well. The cross products are then substituted by a corresponding product of the quaternions, with explicit formulas for the Lie–Poisson bracket closely resembling (3.41). Since the derivation is analogous to Section 3.3, it will be omitted from the exposition.

Appendix D. The subgroup covariant Lagrange–Poincaré approach

In Section 7, we performed Lagrange–Poincaré reduction by considering $SE(3)$ as a manifold with a $SO(3)$ action. In this Appendix, we take into account the full group structure of $SE(3)$ and perform covariant group reduction as in [27] on the charged strand. More precisely, we see the principal bundle

$$\pi_{SE(3)} : SE(3) \rightarrow \mathbb{R}^3, \quad \pi_{SE(3)}(\Lambda, \mathbf{r}) = \Lambda^{-1}\mathbf{r} = \boldsymbol{\rho},$$

as being associated to the left subgroup action of $SO(3) \cong SO(3) \times \{0\}$ on $SE(3)$. Using the composition law in $SE(3)$

$$(\Lambda_1, \mathbf{r}_1)(\Lambda_2, \mathbf{r}_2) = (\Lambda_1\Lambda_2, \mathbf{r}_1 + \Lambda_1\mathbf{r}_2),$$

we obtain that the subgroup action is given by

$$SO(3) \times SE(3) \rightarrow SE(3), \quad \Lambda_1(\Lambda_2, \mathbf{r}_2) = (\Lambda_1\Lambda_2, \Lambda_1\mathbf{r}_2).$$

We thus have recovered the action (7.1). The projection $\pi_{SE(3)}$ identifies an equivalence class $[\Lambda, \mathbf{r}] \in SE(3)/SO(3)$ with the vector $\Lambda^{-1}\mathbf{r} \in \mathbb{R}^3$.

Therefore, we can obtain the equation of the molecular strand by reducing the principal $SE(3)$ -bundle P by the subgroup $SO(3)$. Such a theory, developed in [27], is applied below directly to the n -dimensional generalization of the molecular strand. The difference with the approach described in Section 7 is that here we make concrete use of the group structure of $SE(3)$ and the fact that the principal bundle $\pi_{SE(3)} : SE(3) \rightarrow \mathbb{R}^3$ is associated to a subgroup action.

Recall from Section 2.4 that, for the n -dimensional generalization of the molecular strand, the Euclidean group $SE(3) \simeq SO(3) \ltimes \mathbb{R}^3$ is replaced by an arbitrary semidirect product $\mathcal{O} \ltimes E$ and the spacetime $X = I \times \mathbb{R}$ is replaced by $X = \mathcal{D} \times \mathbb{R}$, where \mathcal{D} is a smooth manifold of dimension n .

Consider the trivial principal S -bundle, $P = X \times S \rightarrow X$, where S is the semidirect product group $\mathcal{O} \ltimes E$. Since P is trivial, the first jet bundle is given by $J^1P_{(x,g)} = L(T_xX, T_gS)$. A section σ of P reads

$$\sigma(x) = (x, \Lambda(x), r(x)), \quad x = (s, t),$$

and its first jet extension is

$$\begin{aligned} j^1\sigma(x) &= (x, \text{id}_{T_xX}, \Lambda(x), r(x), T_x\Lambda, T_xr) \\ &= (x, \text{id}_{T_xX}, \Lambda(x), r(x), \mathbf{d}\Lambda + \dot{\Lambda}dt, \mathbf{d}r + \dot{r}dt), \end{aligned}$$

where \mathbf{d} denotes the derivative with respect to space and the dot denotes the time derivative.

We can see P as an \mathcal{O} -principal bundle over $\Sigma := P/\mathcal{O} = X \times E$, relative to the projection $\iota : (x, \Lambda, r) \mapsto (x, \Lambda^{-1}r)$. Suppose that we have an \mathcal{O} -invariant Lagrangian density \mathcal{L} defined on J^1P . This Lagrangian induces a reduced Lagrangian density $\tilde{\mathcal{L}} : J^1P/\mathcal{O} \rightarrow \mathbb{R}$. On the principal bundle $P \rightarrow P/\mathcal{O}$ we consider the flat principal connection

$$A(v_x, v_\Lambda, v_r) = v_\Lambda \Lambda^{-1}.$$

where $v_x \in T_x X$, $v_\Lambda \in T_\Lambda \mathcal{O}$, and $v_r \in T_r E$. Using this connection, we have the fiber bundle isomorphism $J^1 P/\mathcal{O} \cong J^1(P/\mathcal{O}) \times_\Sigma L(TX, \text{ad } P)$ over Σ . Note that in our particular case, the vector bundle $\text{ad } P \rightarrow \Sigma$ is trivial (see (7.9)) and can be identified with $\Sigma \times \mathfrak{o}$, where \mathfrak{o} is the Lie algebra of \mathcal{O} . Moreover, the connection A is identified with the trivial connection; see (7.10).

We recall now from [27] the covariant Lagrange–Poincaré reduction, adapted here to the case of a semidirect product $S = \mathcal{O} \ltimes E$, and to the fact that P , as a S -principal bundle, is trivial.

Given a section $\sigma = (\Lambda, r)$ of $P \rightarrow X$, we introduce the section σ_1 of $P/\mathcal{O} \rightarrow X$ defined by $\sigma_1(x) := (\iota \circ \sigma)(x) = \Lambda(x)^{-1}r(x) = \rho(x)$, and the section σ_2 of $L(TX, \mathfrak{o}) \rightarrow X$ defined by $\sigma_2(x) = \Lambda^{-1}T_x \Lambda = \Omega + \omega dt$. The following are equivalent:

- σ is a critical point for the variational principle

$$\delta \int_X \mathcal{L}(j^1 \sigma) = 0;$$

- σ satisfies the Euler–Lagrange equations for \mathcal{L} ;
- the variational principle

$$\delta \int_X \bar{l}(j^1 \sigma_1, \sigma_2) = 0$$

holds for arbitrary variations $\delta \sigma_1$ and variations of the form

$$\delta \sigma_2 = \mathbf{d}^A \eta + [\sigma_2, \eta],$$

where η is an arbitrary section of $L(TX, \mathfrak{o}) \rightarrow X$;

- the sections σ_1, σ_2 satisfy the covariant Lagrange–Poincaré equations

$$\begin{cases} \frac{\delta \bar{l}}{\delta \sigma_1} - \text{div}_x \left(\frac{\delta \bar{l}}{\delta (T\sigma_1)} \right) = 0, \\ \text{div}_x^A \frac{\delta \bar{l}}{\delta \sigma_2} = \text{ad}_{\sigma_2}^* \frac{\delta \bar{l}}{\delta \sigma_2}, \end{cases} \quad (\text{D.1})$$

where div_x^A denotes the covariant divergence associated to A and acting on $\mathfrak{X}(X, \mathfrak{o}^*)$. Note that here $\text{div}_x^A = \text{div}_x$.

Using the decomposition $X = \mathcal{D} \times \mathbb{R}$, we can write

$$\mathcal{L}(j^1 \sigma) = \mathcal{L}(\dot{\Lambda}, \mathbf{d}\Lambda, \dot{r}, \mathbf{d}r)$$

and

$$\bar{l}(j^1 \sigma_1, \sigma_2) = \bar{l}(\rho, \Omega, \omega).$$

Hence, we obtain the equality

$$\frac{\delta \bar{l}}{\delta \sigma_2} = \frac{\delta \bar{l}}{\delta \Omega} + \frac{\delta \bar{l}}{\delta \omega} \partial_t.$$

Since $\operatorname{div}_x^A \pi_1^* = \frac{d}{dt}$ and $\operatorname{div}_x^A \pi_2^* = \operatorname{div}$, the second equation of (D.1) reads

$$\frac{d}{dt} \frac{\delta \bar{l}}{\delta \omega} + \operatorname{div} \frac{\delta \bar{l}}{\delta \Omega} = \operatorname{ad}_\omega^* \frac{\delta \bar{l}}{\delta \omega} + \operatorname{ad}_{\Omega_i}^* \frac{\delta \bar{l}}{\delta \Omega_i}.$$

The first equation reads

$$\frac{\delta \bar{l}}{\delta \rho} - \frac{d}{dt} \frac{\delta \bar{l}}{\delta \rho_t} - \operatorname{div} \frac{\delta \bar{l}}{\delta \rho_s} = 0.$$

We have thus obtained equations (6.21) by covariant Lagrange–Poincaré reduction. Of course, when $\mathcal{O} \otimes E = SO(3) \otimes \mathbb{R}^3 = SE(3)$ and $\mathcal{D} = [0, L]$ we recover equations (7.15) and (7.16) for the molecular strand in the new variables.

References

1. ANTMAN, S.S.: *Nonlinear Problems of Elasticity*. Springer, New York, 2004
2. BAKER, N.A., SEPT, D., JOSEPH, S., HOLST, M.J., MCCAMMON, J.A.: Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA* **98**, 10037–10041 (2001)
3. BALAEFF, A., MAHADEVAN, L., SCHULTEN, K.: Elastic rod model of a DNA loop in the lac operon. *Phys. Rev. Lett.* **83**, 4900–4903 (1999)
4. BANAVAR, J.R., HOANG, T.X., MADDOCKS, H.H., MARITAN, A., POLETO, C., STASIAK, A., TROVATO, A.: Structural motifs of macromolecules. *Proc. Natl. Acad. Sci. USA* **104**, 17283–17286 (2007)
5. BLOCH, A.M.: Nonholonomic mechanics and control. In: *Interdisciplinary Applied Mathematics*, vol. 24. Springer, New York (2003). With the collaboration of J. Baillieul, P. Crouch and J. Marsden, With scientific input from P. S. Krishnaprasad, R. M. Murray and D. Zenkov, Systems and Control
6. CENDRA, H., MARSDEN, J.E., RATIU, T.S.: *Lagrangian Reduction by Stages*, vol. 152. *Memoirs American Mathematical Society* (2001)
7. DICHMANN, D., LI, Y., MADDOCKS, J.: Hamiltonian formulations and symmetries in rod mechanics. In: *Mathematical Approaches to Biomolecular Structure and Dynamics* (Minneapolis, MN, 1994), *IMA Volume of Mathematical Applications*, vol. 82, pp. 71–113. Springer, New York, 1996
8. DILL, E.H.: Kirchhoff’s theory of rods. *Arch. Hist. Exact Sci.* **44**, 1–23 (1992)
9. ELLIS, D.C.P.: Reduction by symmetry—an application driven perspective. Thesis, 2010, in preparation
10. ELLIS, D.C.P., GAY-BALMAZ, F., HOLM, D.D., RATIU, T.S.: Lagrange–Poincaré field equations, 2010, preprint
11. GAY-BALMAZ, F., HOLM, D.D., RATIU, T.S.: Variational principles for spin systems and the Kirchhoff rod. *J. Geome. Mech.* **1**(4), 417–444 (2009)
12. GAY-BALMAZ, F., RATIU, T.: The geometric structure of complex fluids. *Adv. Appl. Math.* **42**(2), 176–275 (2008)
13. GAY-BALMAZ, F., RATIU, T.S.: A new Lagrangian dynamic reduction in field theory. *Ann. Inst. Fourier* (2010), to appear
14. GOLDSTEIN, R., GORIELY, A., HUBER, G., WOLGEMUTH, C.: Bistable helices. *Phys. Rev. Lett.* **84**, 1631–1634 (2000)
15. GOLDSTEIN, R., POWERS, T.R., WIGGINS, C.H.: Viscous nonlinear dynamics of twist and writhe. *Phys. Rev. Lett.* **80**, 5232–5235 (1998)
16. GORIELY, A., TABOR, M.: New amplitude equations for thin elastic rods. *Phys. Rev. Lett.* **77**, 3537–3540 (1996)

17. HAUSRATH, A., GORIELY, A.: Repeat protein architectures predicted by a continuum representation of fold space. *Prot. Sci.* **15**, 1–8 (2006)
18. HOLM, D.D.: Euler–Poincaré dynamics of perfect complex fluids. In: *Geometry, Mechanics and Dynamics*, pp. 113–167. Special Volume in Honor of J. E. Marsden, 2001
19. HOLM, D.D.: *Geometric Mechanics Part 2: Rotating, Translating and Rolling*. Imperial College Press, 2008
20. HOLM, D.D., KUPERSHMIT, B.: The analogy between spin glasses and Yang–Mills fluids. *J. Math. Phys.* **29**, 21–30 (1988)
21. HOLM, D.D., MARSDEN, J.E., RATIU, T.: The Euler–Poincaré equations and semidirect products with applications to continuum theories. *Adv. Math.* **137**, 1–81 (1998)
22. HOLM, D.D., MARSDEN, J.E., RATIU, T.S.: *Hamiltonian Structure and Lyapunov Stability for Ideal Continuum Dynamics*. University of Montreal Press, Montreal, 1986
23. HOLM, D.D., PUTKARADZE, V.: Nonlocal orientation-dependent dynamics of molecular strands. *C. R. Acad. Sci. Paris* **347**, 1093–1098 (2009)
24. KEHRBAUM, S., MADDOCKS, J.H.: Elastic rods, rigid bodies, quaternions, and the last quadrature. *Phil. Trans* **355**, 2117–2136 (1997)
25. KIRCHHOFF, G.: Über das Gleichgewicht und die Bewegung eines unendlich dünnen elastischen Stabes. *J. f. reine angew. Math.* **56**, 285–313 (1859)
26. KRIEGL, A., MICHOR, P.W.: The convenient setting of global analysis. In: *Mathematical Surveys and Monographs*, vol. 53. American Mathematical Society, Providence, 1997
27. LÓPEZ, M.C., RATIU, T.S.: Reduction in principal bundles: Covariant Lagrange–Poincaré equations. *Commun. Math. Phys.* **236**, 223–250 (2003)
28. MARSDEN, J.E., MISIOLEK, G., ORTEGA, J.P., PERLMUTTER, M., RATIU, T.S.: Hamiltonian Reduction by Stages. In: *Lecture Notes in Mathematics*, vol. 1913. Springer, Berlin, 2007
29. MARSDEN, J.E., RATIU, T., WEINSTEIN, A.: Semidirect products and reduction in mechanics. *Trans. Am. Math. Soc.* **281**, 147–177 (1984)
30. MARSDEN, J.E., RATIU, T.S.: Introduction to mechanics and symmetry: a basic exposition of classical mechanical systems. In: *Texts in Applied Mathematics*, 2nd edn. Springer, Berlin, 2002
31. MEZIC, I.: On the dynamics of molecular conformation. *Proc. Natl. Acad. Sci. USA* **103**, 7542–7547 (2006)
32. MOAKHER, M., MADDOCKS, J.H.: A double-strand elastic rod theory. *Arch. Rational Mech. Anal.* **177**, 53–91 (2005)
33. MORITA, H., KANEKO, K.: Self-organized bottleneck in energy relaxation. *Europhys. Lett.* **66**, 198–204 (2004)
34. NEUKIRCH, S., GORIELY, A., HAUSRATH, A.C.: Chirality of coiled coils: Elasticity matters. *Phys. Rev. Lett.* **100**, 038,105 (2008)
35. NIZETTE, M., GORIELY, A.: Towards a classification of Euler–Kirchhoff filaments. *J. Math. Phys.* **40**, 2830–2866 (1999)
36. NODA, K., ISHIDA, K., KUBONO, A., HORIUCHI, T., YAMADA, H.: Remanent polarization of evaporated films of vinylidene fluoride oligomers. *J. Appl. Phys.* **93**, 2866–2870 (2003)
37. SIMÓ, J.C., MARSDEN, J.E., KRISHNAPRASAD, P.S.: The Hamiltonian structure of non-linear elasticity: the material and convective representations of solids, rods, and plates. *Arch. Rat. Mech. Anal.* **104**, 125–183 (1988)

Mathematics Department, Imperial College London,
London SW7 2AZ, UK.
e-mail: de102@imperial.ac.uk

and

Laboratoire de Météorologie Dynamique,
École Normale Supérieure/CNRS, Paris, France.

and

Control and Dynamical Systems,
California Institute of Technology,
Pasadena, CA 91125, USA.
e-mail: fgbalmaz@cds.caltech.edu

and

Institute for Mathematical Sciences,
Imperial College London,
London SW7 2PG, UK.

and

Mathematics Department,
Imperial College London,
London SW7 2AZ, UK.
e-mail: d.holm@imperial.ac.uk

and

Department of Mechanical Engineering,
University of New Mexico,
Albuquerque, NM 87131-1141, USA.

and

Department of Mathematics,
Colorado State University,
Fort Collins, CO 80523-1874, USA.
e-mail: putkarad@math.colostate.edu

and

Section de Mathématiques and Bernoulli Center,
École Polytechnique Fédérale de Lausanne,
1015 Lausanne, Switzerland.
e-mail: Tudor.Ratiu@epfl.ch

(Received February 17, 2009 / Accepted October 25, 2009)
Published online April 28, 2010 – © Springer-Verlag (2010)